PREVALENCE AND INCIDENCE OF HORMONAL-RELATED CANCERS (HRCs) IN RURAL SASKATCHEWAN:

THE SASKATCHEWAN RURAL HEALTH STUDY

A Thesis Submitted to the College of

Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Community Health and Epidemiology

University of Saskatchewan

Saskatoon

By

Amal S. Khan

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Request for permission to copy or to make other use of the material in this thesis in whole or part should be addressed to:

Head of the Department of Community Health and Epidemiology

The University of Saskatchewan.

Saskatoon, Saskatchewan (S7N-5E5).

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9 Canada

ABSTRACT

A significant number of rural Canadians dwellers are affected by hormone-related cancers (HRCs) like breast and prostate cancer. Little is known about the rates of occurrence and contextual factors that are associated with the development of breast and prostate cancers in rural residents. This study aimed to determine the incidence and prevalence of breast and prostate cancer among Saskatchewan rural dwellers and to explore individual and contextual factors that are associated with the prevalence of these cancers. To accomplish our goal, we utilized and analyzed data from the Saskatchewan Rural Health Study (SRHS).

The SRHS involved a prospective cohort conducted in two phases: the baseline survey (in 2010) and a 4-year follow-up survey (in 2014). In the baseline, the SRHS research team obtained completed questionnaires from 4624 households, including information about 8261 individuals, 18 years and older. Questionnaires were returned from 2797 households comprised of 4867 individuals, in the follow-up survey.

Crude prevalence and incidence of HRCs were calculated using appropriate formulae. Adjusted prevalence was calculated using logistic regression based on generalized estimating equations approach to account for hierarchy in the data (individuals within a household).

Our study reported the crude prevalence of HRCs (breast and prostate cancers combined) as 3.0%, and 3.4% respectively, at the baseline and follow-up. The adjusted prevalence analysis showed that following variables were significantly associated with HRCs (breast and prostate cancers combined): exposure to radiation (odds ratio [OR] 3.39; 95% confidence interval [CI]: 2.23, 4.84), previous history of cancer in a sibling (i.e. brother or sister) (OR = 1.51, 95% CI: 1.11, 2.07) and a positive history of cancer in father (OR=1.37; 95% CI= 1.01, 1.86).

The current study showed cumulative incidence for breast and prostate cancers in rural Saskatchewan as 0.86% and 1.08%, respectively. When combined (i.e. breast and prostate cancers together), the cumulative incidence for HRCs was 0.97%.

In summarizing our study findings, HRCs were slightly more prevalent amongst non-farm residents 3.2% when compared with farm residents 2.8% and within the eastern part of the province (6.5%, 6.6%, respectively) as compared to the western part (4.6%, 6.3%, respectively) among both farm and non-farm residents, at the baseline. It appears that the prevalence of HRCs among farm and non-farm rural residents depend on the complex interplay among a variety of factors such as individual and contextual factors.

ACKNOWLEDGMENTS

All praises belong to God alone, who is the creator of (all the infinities of) the heavens and the earth to whom I belong and to whom all thanks are due. While the world looks upon me as I struggle along, God (alone) showed himself considerate and gentle towards me as if I have not sinned, therefore, O my lord I find thee the most praiseworthy and thou (alone) deserve and worthy of the tribute I pay to thee.

I would like to extend a note of deep gratitude towards my supervisors Dr. Punam Pahwa and Dr. Shahid Ahmed for helping me immensely in achieving the milestones during my journey at the graduate school. No words can express my gratitude to both of you for your precious and sincere support that enabled me to complete my thesis successfully. I especially thank Dr. Pahwa for her advice and to Dr. Ahmed, who has always been just an email away.

I am very grateful to my committee members; Dr. Rachel and Dr. Adnan Zaidi for their valuable feedback on my research work. Dr. Rachel endorsed me worthy suggestions on the epidemiological concepts of this research project and proved to be a great support in refining my writing skills. Dr. Zaidi was undoubtedly a source of invaluable guidance on my thesis that gave this research project a complete finish.

A special thanks to Dr. James Dosman, the nominated principal investigator of the CIHR-funded Saskatchewan Rural Health Study. I would like to extend my indebtedness gratitude to Dr. Dosman for providing me the data which I have used in accomplishing this research project.

Dr. Anurag Saxena and Dr. Andrew Lyon have always been a great support and source of encouragement for me. I am grateful to them for their kind gesture and moral support. Their confidence in my abilities and appreciation enabled me to steadfast in achieving my goal.

Lastly, I would like to express my deepest gratitude to my parents, my husband Muhammad, and especially to my little ones; Mustafa and Muaz who sacrificed on my behalf. I owe them my success, and I acknowledge them to be the true constant companions throughout.

DEDICATION

This work is dedicated to my mama Tahseen S. Khan and dady Muhammad S. Khan for their prayers, love, inspiration, dedication, and endless efforts towards bringing me to this stage. Dady, all this could not have been accomplished if you were not beside me. I do not know what I will do without you, you are truly my superhero, and I love you so much! I love you more than the stars in the sky, more than the fish in the sea, more than the tree in the forest. Dady you mean the world to me!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

HRCs – Hormone Related Cancers

SRHS – Saskatchewan Rural Health Study

WHO – World Health Organization

AICR – American Institute of Cancer Research

GLOBOCAN - Global Cancer Statistics Canada

CMA - Canadian Medical Association

CNBSS - Canadian National Breast Screening Study

SES – Socio-Economic Status

RR - Relative Risk

CI – Confidence Interval

SRR – Standardised Relative Risk

WCRF - World Cancer Research Fund

RMs – Rural Municipalities

PHF - Population Health Framework

BMI – Body Mass Index

LDA - Longitudinal Data Analysis

GLMs - Generalized Linear Models

GEE – Generalized Estimating Equation

ORs – Odds Ratio, aOR - Adjusted Odds Ratio

PH - Cox Proportional Hazard

QIC - Quasi-likelihood under the Independence model Criterion

OLR - Ordinary Least square Regression

FDRs - First-Degree Relatives

HR – Hazards Ratio

CCS – Canadian Cancer Society

CCS - Canadian Cancer Statistics

SCR - Saskatchewan Cancer Registry

SIR – Standardized Incidence Ratio

PSA - Prostate Specific Antigen

CHAPTER 1: INTRODUCTION & RATIONALE

1.1 Introduction

Cancer denotes a disease which is characterized by the formation of abnormal cells that divide in an uncontrollable fashion, can infiltrate, and destroy normal body cells (1). According to recent data by the World Health Organization (WHO), cancer is the leading cause of death worldwide and accounts for around 22% of all deaths not related to communicable diseases (2). Globally, deaths from cancer are projected to continue rising with an estimation of about 70% increase by 2030 (2). A considerable number of cancer-deaths are attributed to hormone-related cancers (i.e. breast cancer mainly) (2).

Hormone-related cancers (HRCs) which include breast, prostate, ovary, testis, endometrium, thyroid, and osteosarcoma (3), as well as well-differentiated neuroendocrine tumor such as insulinomas, are believed to share a unique mechanism of carcinogenesis. HRCs involve cell proliferation by endogenous and exogenous hormones, which cause cell division and increase the chance of random genetic error for the initiation of cancer (3). Breast and prostate cancers are the two major cancers within HRCs. It is against this backdrop that this thesis focuses on these two major HRCs. From here-on in this thesis, the term HRCs synonymously refers to breast and prostate cancers.

Globally, breast cancer is the most commonly diagnosed cancer among women in 140 countries and is the most common cause of cancer death in women in 103 countries worldwide (4). American Institute of Cancer Research (AICR) declared prostate cancer as the second most common cancer in men worldwide (5). In North America, prostate cancer (excluding skin cancers) remained the most commonly diagnosed cancer among men and the second most common cause of death (6). Similarly, in women, breast cancer has been identified as the most commonly diagnosed cancer, and the second most common cause of cancer death among them (6).

In Canada, malignancy of breast accounts for approximately (25%) of all new cases of cancer among females (122). For Canadian males, prostate cancer attributes to around one-fifth (20%) of all new cancer cases (122).

Despite higher rates of mortality and morbidity associated with HRCs, the etiology of such cancers remains inconsistently reported. Literature has highlighted some of the putative risk factors

associated with HRCs that include; age (7, 9), family history (8-9), geographical location (10-11), socioeconomic status (12-13), diet (14-15), physical activity (16-17), environmental and occupational exposures (8, 18).

In terms of age as a risk factor for breast cancer, it has been established that the incidence rate increases with advancing age (19) even though recent data show a diminishing risk of breast cancer occurrence at older ages as compared to younger ages (20). Similarly, the rate of prostate cancer incidence increases with age (9) to about 984.8/100,000 among men of age $\geq$70 years after which it slightly decreases (21). For men aged 40–44 years, the incidence rate of prostate cancer is 9.2/100,000 (21).

In an epidemiological review of breast cancer risk factors, dietary behavior was reported to influence both the promotion as well as inhibition of breast cancer development (22). Studies exploring the causal pathways of dietary effects on prostate cancer have reported inconsistent findings (23).

Due to such inconsistencies regarding the risk factors associated with HRCs, there is a need for further studies to explore and understand the potential risk factors. Also, the prevalence and incidence of HRCs appear not have been studied adequately in Canada, and in particular, among rural populations.

We, therefore, propose to address the aforementioned gaps by investigating the prevalence and incidence of HRCs in rural Saskatchewan using data from the Saskatchewan Rural Health Study (SRHS) which surveyed rural dwellers from across Saskatchewan. As a result, we will investigate the association between the risk factors associated with the prevalence of HRCs among farm and non-farm resident women and men, respectively, in rural Saskatchewan.

1.2 Rationale

Rural Canada accounts for about 90% of the country's total landmass and is home to a significant proportion of Canada's population (24). Recent data show that the proportion of the rural population in Saskatchewan (33%) is almost twice the national proportion (18%) (25-26). Also, it has been reported that about 3 to 6% of global cancers are caused by exposure to carcinogenic substances at workplaces (25, 27).

Home to the largest rural population in Canada, the major occupation of rural residents in Saskatchewan is farming (26, 28). Interestingly, among all cancers, breast and prostate cancers have been reported to be the most prevalent cancer-types in Saskatchewan (29). Canadian Cancer

Statistics (CCS) 2019 report identified projected age-standardized incidence rates (ASIR) for selected cancers, by sex and province, in Canada (excluding Quebec) (123).

CCS reported ASIR for prostate cancer as (117.8 per 100,000) out of all cancers in males (533.1 per 100,000) and ASIR for breast cancer as (122.9 per 100,000) out of all cancers (452.7 per 100,000) in females (123). These reported ASIRs indicate that among all cancers in females, breast malignancy accounts for around (27.15%) and for males, prostate malignancy accounts for about (22.1%) of all male malignancies, in the province of Saskatchewan (123). In the same report, the rates were age-standardized to the 2011 Canadian population (123).

This delineates the extent of the burden of breast and prostate cancers in the province of Saskatchewan, where the rural population, who are predominantly farmers, constitutes a significant proportion of the population. With such a huge disease burden, there still appears to be a paucity of information about the prevalence and incidence of breast and prostate cancers in Saskatchewan.

In fact, to the best of my search into literature, I could not find any study exploring the prevalence and incidence of HRCs in rural Saskatchewan. The recent trends in HRCs (as discussed above), combined with the limited availability of studies on such cancers in Saskatchewan calls for more research to explore the prevalence and incidence of breast and prostate cancers as well as the risk factors associated with these risk measures.

Studying the prevalence of breast and prostate cancers will help measure the pattern of pervasiveness of such cancers in this Saskatchewan rural cohort. The measure of the incidence of breast and prostate cancers will help us delineate the frequency of occurrence of HRCs. This will guide us towards exploring reasons for such a probable trend in rural Saskatchewan and help directly in influencing healthcare resource allocation within the province of Saskatchewan and Canada at large.

In addition, identifying the potential risk factors associated with breast and prostate cancers will guide investigation to the underlying etiological factors that contribute to the prevalence of these cancers. Establishing the risk factors of breast and prostate cancers will also aid in targeting future prevention and may influence treatment options for people diagnosed with these cancers.

1.3 Objectives

Overall objective: The overall objective is to determine the prevalence, incidence, and risk factors associated with the prevalence of hormone-related cancers (i.e., breast and prostate cancer) among Saskatchewan rural dwellers 18 years and older.

Research questions:

1. What is the prevalence of breast and prostate cancers among women and men, respectively in rural Saskatchewan?
2. What is the incidence of breast and prostate cancer among women and men, respectively in rural Saskatchewan?
3. What are the risk factors associated with the prevalence of breast and prostate cancers among women and men, respectively in rural Saskatchewan?

CHAPTER 2: LITERATURE REVIEW

2.0 Literature review

In this chapter, I present a review of relevant literature for the purposes of this thesis. The remainder of this chapter is organized as follows: Section 2.1 provides a brief background while the oncogenesis/carcinogenesis is provided in Section 2.2. A discussion of HRCs is provided in Section 2.3 and the epidemiology of HRCs at the international, national, and local levels is provided in Section 2.4. Finally, a distinction between the etiology and risk factors of HRCs is provided in Section 2.5.

2.1 Background

As organs, breast and prostate are different in terms of their respective anatomical and physiological functions (30). However, both organs depend on gonadal steroids for their development and the neoplasms that grow in them are typically hormone-dependent and have extraordinarily underlying overlapping biology (30). Since the 1950s, epidemiological studies investigating the genetic and biological aspects of breast and prostate cancers have reported a higher frequency of prostate cancer among the relatives of women with breast malignancy (31).

A similar study investigated the genetic bases of breast cancer and propose that "prostate cancer could be the male equivalent of at least some female mammary carcinomas" (31). A review of the frequency of breast and prostate cancers (between the 1960s – 1980s) in 21 countries showed a high correlation between the rates of incidence of both tumors (32). A similar correlation was reported between the diagnosis of breast and prostate cancers and endogenous sensitivities as well as constitutional factors including metabolic, genetic, hormonal, and environmental risk factors (32).

2.2 Oncogenesis/Carcinogenesis

Oncogenesis, also called carcinogenesis or tumorigenesis, refers to the process in which normal body cells are transformed into cancerous cells via abnormal cell division (33). Usually, in the human body, a balance between the proliferation of cells and programmed cell death is maintained, and disturbances in this equilibrium often arise through mutations and epimutations (33). Oncogene denotes a gene that plays a normal role in the cell as a proto-oncogene (34).

Mutational alteration in proto-oncogene may cause the growth of a tumor (34). Some important oncogenes include ras, Bcl-2, myc, HER-2/neu, and hTERT (34). Mutations in the genes may be

inherited (genetic predisposition) or acquired (i.e. caused by exposure to carcinogenic substances in the environment) or a combination of both (35). Moreover, carcinogenesis can result from a combination of biological, genetics, chemical and a physical insult to human cells (36).

2.3 Hormone Related Cancers

Epidemiological studies have established that hormone-related cancers share a unique mechanism of carcinogenesis (3). Hormones influence the malignancies of the breast, prostate, endometrium, ovary, testis cancers (37), etc. Although hormones (i.e. endogenous, exogenous) contribute to cell proliferation which gives rise to random genetic errors, yet the emergence of a malignant phenotype depends on a series of somatic mutations during cell division (3).

In addition, the specific genes that cause the progression of such cancers (HRCs) remain ill-defined (3). Also, carcinogenesis in hormone-related cancers is an entirely different model which itself details the importance of cell proliferation in the targeted population of a cell (38). Thus, the formation of cancer in glands like mammary glands and prostate is complex and multifactorial and may result from a variety of factors including biological, non- biological, and social factors (i.e. hereditary, endocrinologic factors, viruses, chemicals, ionizing radiations, diet, socio-economical, familial factors, etc.) (39).

In particular, hormones (i.e. estrogen and progesterone) that are considered as the two main female sex hormones (127) promote malignancy of breast (123). This is achieved by the proliferation of cancer cells through induction (123) of growth-promoting proto-oncogenes (124) encoding for cyclin-G1 expression in estrogen- and progesterone-mediated breast cancer (123).

Androgens (i.e. testosterone and androstenedione) are produced by both males and females' bodies and are linked to play a significant role in males' trait and reproductive activity (125) by regulating the physiological as well as the pathological developments associated with prostatic disease (126). Although the mechanisms that androgens utilize to influence these pathways remains to be not well understood (126).

2.4 Epidemiology of Hormone Related Cancers (HRCs)

In this section, international, national, and local trends in the epidemiology of HRCs are highlighted. In particular, Canadian and Saskatchewan trends are presented.

2.4.1 International prevalence and incidence

According to a 2012 GLOBOCAN (Global Cancer Statistics Canada, 2012) report, approximately 14.1 million new cases of cancers and 8.2 million cancer-related deaths occurred worldwide (40).

Of these, breast cancer is the most frequently diagnosed cancer and is considered the leading cause of cancer death in women in less-developed countries (40). Of note, cancer of the female breast has been regarded as the most common cancer of the developed countries as well, with every one in ten of all new cancer cases diagnosed worldwide each year (41). The same study called breast cancer-related mortality as the principal cause of death from cancer among women globally (41).

Similarly, carcinoma of prostate stands as the second most commonly diagnosed malignancy in males worldwide (42). In developed countries, the incidence of prostate cancer is the most common in terms of its diagnosis among men (40). In 2012, approximately 1.1 million men were diagnosed with carcinoma prostate globally, constituting about 15% of the cancers diagnosed in men (43). Out of this, almost 70% (759,000) occurred in more developed areas (43). Moreover, crude differences in prostate cancer incidence between developed and developing regions are probably due to underutilization of screening (44) along with lower life expectancies in developing regions of the world (45).

Generally, global estimates of cancer prevalence for 27 sites in an adult population indicate that breast cancer continues to be the most prevalent cancer in the vast majority of countries (46). As it has been estimated that, with 1 million new cases in the world annually, breast cancer continues to be the commonest malignancy in women comprising 18% of all female cancers worldwide (47).

On the other hand, carcinoma of prostate stands as the sixth common malignancy worldwide, the second most common cancer in males, and the commonest one in men in parts of Africa, Europe, and North America (48, 49). Carcinoma of the prostate will be regarded as the most common malignancy in men in the future (50). This could be due to an increase in the screening practices of prostate-specific antigen (PSA) in different regions of the world (50).

2.4.2 Canadian prevalence and incidence

According to the Canadian Cancer Society (CCS), cancer is the leading cause of death and constitutes about 30% of all deaths in Canada (51). The probability of developing cancer is 1 in every 2 Canadians (49% of men & 45% of women) and 1 in every four Canadians (28% of men & 24% of women) will die from the same (51).

The incidence and prevalence of cancer cases will continue to increase due to a sustained rise in the number of cases diagnosed each year and improvement in the survival rates (52). This could

partly be attributed to advancement in screening techniques, improvement in treatment, and, cancer palliative care (45).

Data from the Canadian Medical Association (CMA) showed that prostate cancer was the most diagnosed form of cancer in 2012 (26 500 new cases or 121 per 100 000 men) (53).

In terms of breast cancer, a randomized screening trial in 2014 by the Canadian National Breast Screening Study (CNBSS) showed that 6383 (7.1%) of 89 835 women developed breast cancer after 25 years of follow-up (54).

A study conducted in 2005 measuring cancer prevalence among the Canadian population reported breast cancer (20.6% of ten-year prevalent cases) and prostate cancer (18.7%) as the most prevalent (55). Breast and prostate cancers accounted for the highest ten-year prevalence cases among women (40.0%) and men (38.2%), respectively (55).

2.4.3 Saskatchewan prevalence and incidence

Literature regarding the prevalence and incidence of HRCs was reviewed. Recent cancer statistics from Saskatchewan highlighted breast and prostate cancers as the most prevalent cancer-types in Saskatchewan (29). In terms of cancer incidence, Canadian Cancer Statistics (CCS) in 2017 revealed that breast and prostate cancers were the most frequently diagnosed type of cancer among Saskatchewan women and men, respectively (56).

This was further supported by Saskatchewan cancer statistics at glance in the Canadian Cancer Statistics 2019 report, which stated that, for men in Saskatchewan, prostate cancer is the most frequently diagnosed type of cancer (57) and for Saskatchewan women, breast cancer is the most frequently diagnosed type of cancer (57).

In addition, the Canadian Cancer Society (CCS) 2019 report identified age-standardized incidence rates (ASIR) for selected cancers, by sex and province, in Canada (excluding Quebec) (123). This report (123) indicated ASIR for prostate cancer as (117.8 per 10^5) out of all cancers in males (533.1 per 10^5) and ASIR for breast cancer as (122.9 per 10^5) out of all cancers (452.7 per 10^5) in females (123).

The above-reported ASIRs indicate that among all cancers in females, breast malignancy accounts for around (27.15%) and for males, prostate malignancy accounts for about (22.1%) of all male malignancies, in Saskatchewan (123).

The above statistics highlight the burden of HRCs in the province of Saskatchewan.

2.5 Etiological factors Vs. Risk factors

Etiological or aetiological factors are the factors that cause or contribute to the cause of a disease or condition (58). On the other hand, a risk factor is any exposure, attribute or character of an individual that increases the likelihood of developing a disease or condition (59).

Section 2.5.1 discusses the risk factors for HRCs in general. In particular, a discussion on specific risk factors is provided under sections 2.5.2, 2.5.3 which entails a discussion on age, and family history. Geographical locations and SES (socioeconomic status) are discussed in section 2.5.4. Other risk factors i.e. diet, and physical activity are discussed under sections 2.5.5. Finally, section 2.5.6 provides a discussion on the relationship between HRCs and environmental/occupational exposures.

2.5.1 Risk factors for HRCs

Risk factors associated with HRCs have been widely reported in the literature. For this thesis, literature describing the association between HRCs and contextual/environmental, individual, occupational exposures, and important covariates were reviewed. It is established that although both intrinsic and extrinsic risk factors play a role in cancer development through the accumulation of random errors (driver mutations), yet, collectively, cancer risk is heavily influenced by extrinsic factors (60).

Moreover, it has been identified that overall the most common risk factors for cancers include age and family history of cancer, smoking and alcohol, exposures to sun, radiations, and various chemicals (i.e., pesticides), certain hormones (i.e., sex hormones) and lack of physical activities (or being overweight) along with poor diet (61). For HRCs such as breast and prostate cancers, it appears that the disease is multifactorial in etiology, and hence, the specific etiological factors remain unclear in the literature.

However, to date, the well-established factors for prostate cancer include; race, family history, hormones, aging, diet, environmental agents, occupational exposures and lifestyle factors (62). Similarly, the established and probable risk factors for breast cancer are; age & family history, geographical location & socioeconomic status, cancer in the other breast or previous benign disease, diet and body weight, and exposure to radiations, hormones, alcohol, tobacco (62).

Although the risk factors above have been well determined for prostate and breast cancers, however, literature urges for more and in-depth research in scrutinizing the relationship of specific factors that play roles at macro, mesa, and micro levels in the development of HRCs.

2.5.2 Age

Age plays a significant role in breast cancer development as the risk varies with a change in age from menarche to age at first full pregnancy and menopause (47). The incidence of breast cancer increases with age which doubles about every ten years until menopause (47). That is, women who start menstruating early in life or who have late menopause have an increased risk of developing breast cancer [47]. Women who have natural menopause after the age of 55 are twice as likely to develop breast cancer as compared to women who experience menopause before the age of 45 (47).

Age plays a significant role in causing prostate cancer (63). Prostate cancer is rare in men younger than 40 years, and the chance of developing prostate cancer increases rapidly after age 50 (64). Prostate cancer develops typically in older men and about 6 of every 10 cases of diagnosed prostate cancer is reported in men aged 65 or older (64). The average age at diagnosis of prostate cancer is reported to be 66 years (64).

2.5.3 Family history

The relationship between family history of cancer and HRCs have been inconsistently reported in the literature. Prostate cancers have been reported in men without a positive family history of the disease while this malignancy appeared to run in some families suggesting an inherited component (64). Women with close relatives (i.e. sister, mother, and daughter) who had been diagnosed with breast cancer have a higher risk of developing the disease as compared to their normal counterparts (65). For example, having one first-degree female relative (i.e. sister, mother, daughter) diagnosed with breast cancer doubles the risk, and with two first-degree relatives diagnosed with breast cancer, the risk is five times higher than average (65). In some cases, a strong family history of breast cancer has been linked to possessing an abnormal gene (such as the *BRCA1* or *BRCA2* gene) which is associated with a higher risk of developing breast and prostate cancers (65). In other cases, an abnormal *CHEK2* gene may play a role in developing breast cancer (65) and it has also been suggested to increase the risk of prostate cancer (128).

2.5.4 Geographical locations and Socioeconomic status

Utilizing data from Statistics Canada, one of the national posts on Cancer Nation depicted that "*Cancer is an indiscriminate disease, affecting rich and poor, old and young, often for no discernible reason. Still, Canadians' odds of getting sick or dying depend surprisingly on where they live" (111).* Saskatchewan has one of the largest populations of rural residents as compared

to all other provinces in Canada (66). In Canada, residential status (i.e., urban versus. rural) has been recognized as a social determinant of health (67). This is also true when trends of some of the cancer types (e.g., breast cancer) or sub-types (e.g., inflammatory breast cancers) are examined (67).

Moreover, within the same country, cancer rates for some cancer types have been reported higher among the population living in rural areas as compared to those living in urban areas (68). This is true partly due to a wealth of resources, screening programs and services available for people of urban areas than in rural and northern areas (69). A study conducting a consultation with Canadian rural women with breast cancer concluded that rural women have greater challenges in terms of issues of access to cancer-related information, support, and services (69).

It is noteworthy that a metanalysis carried out in 2014 showed, women living in rural areas of Canada were more likely to be diagnosed at an advanced stage of their breast cancer disease compared to the women of urban Canada (70). Perhaps, this contrast reflects well-adopted breast cancer screening programs and practices in the urban regions of the country (69).

Similarly, for prostate cancer, one of the most notable characteristics is the degree of geographical variation in the patterns of its incidence and prevalence which is evident at local, national, and international levels (71). Prostate cancer is the most common in North America, North-Western Europe, the Caribbean island, and Australia compared to Asia, Africa, South and Central America (72). The reason for this high prevalence remains not very clear (72), however, this could partly be attributed to a prostate-specific antigen (PSA) screening practices in the Western world.

In cancer epidemiology, the impact of socioeconomic status (SES) on incidence rate and prognosis has been of concern to researchers. In terms of socioeconomic status and the development of HRCs, a higher level of socioeconomic status (SES) was found to be significantly associated with an increased risk of prostate cancer disease (relative risk (RR) = 1.28); 95% confidence interval (CI): 1.25–1.30] (73). However, for prostate cancer-related deaths, higher levels of SES were found to be significantly related to lower rates of mortality [RR = 0.88; 95% CI: 0.92–0.94] (73).

Besides, women who belong to a higher socioeconomic class have a significantly higher breast cancer incidence (SRR 1.25, 1.17-1.32), which could be explained by factors such as mammography screening, hormone replacement therapy, reproductive and lifestyle factors (74).

However, a decreased case fatality rate (SRR 0.72, 0.63-0.81) has been reported among women from higher socioeconomic status (74).

This is partly explained by differences in treatment factors, lifestyle factors, associated comorbidities, and tumor characteristics (74). A similar study expressed a significantly increased breast cancer-related mortality (SRR 1.16, 1.10-1.23) for women with higher socioeconomic status (74).

2.5.5 Diet and Physical Activity

The relationship between HRCs and diet and/or physical activity has been widely reported in the literature. The World Cancer Research Fund (WCRF) estimates that approximately 20% of all cancers diagnosed in the US were related to practicing sedentary lifestyles such as physical inactivity, excess alcohol consumption, and/or poor nutrition (75). The Canadian Medical Association (CMA) highlighted healthy eating and physical activity as the two most important behaviors to create and maintain good health (76). This is particularly meaningful when one has prostate cancer since being overweight raises the risk of aggressive or advanced prostate cancer (63). Also, physical activity has been increasingly associated with a decreased risk of breast cancer occurrence in both premenopausal and postmenopausal women (77). A meta-analysis of 31 prospective studies reported that work-related, as well as non-work-related physical activities, had a decreased incidence of breast cancer (78). The same report suggested a linear relationship between risk of breast cancer and physical activity (i.e. for every 25 MET-h/week increase in activity the risk of breast cancer decreases by 2%, and for 10 MET-h/week and 2 h/week are 3% and 5%, respectively (78).

In terms of the relationship between dietary practices and HRCs, increase intake of fat from meat has been linked to an increased incidence of prostate cancer in the western industrial world (79). A study analyzing the current human-based evidence revealed that there is limited evidence to support a link between plant-based dietary products with a lower risk of prostate cancer (80).

For breast cancers, it has been reported in the literature that diet plays an important role in promoting as well as inhibiting the development of human breast cancer (81). Regarding the association of dietary fat with breast cancer risk, epidemiological studies have been consistently reporting conflicting results (81).

Overall, since dietary exposure is ubiquitous, and much of the epidemiologic evidence is observational, one of the limitations is that causality is difficult to establish.

2.5.6 Relationship between HRCs and Environmental/Occupational Exposures

The relationship between HRCs and environmental as well as occupational exposures recently appears to be gaining prominence. A study conducted during the 1960, 1970, 1980-81, and/or 1990 census examining the relationship between occupation and cancer covering 15 million people in five Nordic countries (i.e. Sweden, Denmark, Finland, Iceland, and Norway) highlighted this association and called such cancers (cancers that develop as a result of a particular exposure at workplace) as 'Occupational cancers' (82). The same study reported that for all cancers (excluding cutaneous cancer), a wide variation was observed among males from a standardized incidence ratio (SIR) of 1.48 (1.43–1.54) in waiters to 0.79 (95% confidence interval 0.66–0.95) in domestic assistants (82). Among females, the standardized incidence ratios varied from 1.27 (1.19–1.35) in tobacco workers to 0.58 (0.37–0.87) in seafarers (people who regularly travel by sea) (82). Higher SIRs were observed among workers producing tobacco and beverage, seamen and chimney sweeps, whereas low SIRs were found for farmers, gardeners, and teachers (82).

This emphasizes the need to consider occupational exposures as potential risk factors of HRCs among rural Saskatchewan residents whose dominant occupation is farming.

Moreover, it was suggested that environmental factors (related to early and later life events) were important in the etiology of breast cancer and could also have considerable influence on the probability of developing clinically detectable prostate malignancy (83).

A case-control study examining the relation of prostate malignancy with farming highlighted that farming was associated with increased risk of prostate cancer among Caucasians (adjusted odds ratio [aOR] 1.8; 95% confidence interval [CI] 1.3 to 2.7) as compared to African Americans (aOR 1.0; 95% CI 0.6 to 1.6) (84).

The same study reported a higher risk of prostate cancer (aOR 1.6; 95% CI 1.2 to 2.2) among farmers who dealt with pesticides (i.e. applied or mixed pesticides and related substances) (84). The study concluded that racial variation in the relationship between prostate malignancy and farming might be attributed to a difference in gene and environmental interactions (84).

Therefore, epidemiologic studies particularly examining rural environmental and occupational exposures among farmers may help explain the rising trend of such cancers (i.e. breast and prostate cancer) (85).

Analytical studies primarily aimed at identifying carcinogenic exposures in the agricultural and environmental settings are required to understand the rising incidence rates for the respective

tumor in the general population (85). One such study reported a higher risk of developing certain cancers such as Non-Hodgkin lymphoma among Saskatchewan farmers (86). The same study emphasized the need for further studies to obtain a better description of farming-related exposures, and to study the exposure history of individuals who develop any type of cancers (86).

Due to the above-illustrated inconsistencies in the literature regarding the role of occupational and environmental exposures as risk factors for HRCs, further research is needed to help explain these inconsistencies.

As a result, I utilized data from the Saskatchewan Rural Health Study (SRHS) to help understand the prevalence, incidence (crude), and risk factors associated with the prevalence of HRCs among farm and non-farm rural residents of Saskatchewan. The methods used to achieve this goal are discussed in the following sections.

CHAPTER 3: DATA DESCRIPTION & METHODS

This chapter presents a description of the data used in this thesis as well as the statistical methods used to analyze the same. In particular, the first half describes the Saskatchewan Rural Health Study (SRHS) in Section 3.1. Details pertinent to study design is given in section 3.1.1. The study population and methods employed to obtain study samples along with data collection tools are described in sections 3.1.2. The sample size is illustrated under section 3.1.3. The theoretical framework is discussed in section 3.1.4. Details of variables used for the purposes of this thesis are provided in section 3.1.5 and its various sub-sections. Variables reported in this chapter as used in this thesis were all taken from reference number (66) [Pahwa et al (2012)].

The remaining half of this chapter discusses the statistical methodology and application of these methods to answer the objectives contained in this thesis in Section 3.2.

3.1 Data Description

3.1.1 Design of the Study

Data from the Saskatchewan Rural Health Study (SRHS) was used for this thesis project. The SRHS, a prospective cohort study of rural residents in Saskatchewan, Canada, was carried out in two phases: a baseline survey in 2010 and a four-year follow-up in 2014 (87). Before conducting the baseline survey, the administration and content of the baseline questionnaire were piloted and optimized in a pilot study (88).

Responses from this pilot study resulted in the modifications of some questions before the actual administration of the baseline questionnaire to study participants. A similar questionnaire was used to collect data in the follow-up study. The baseline survey was carried out in three distinct stages. Stage one involved partitioning the southern half of Saskatchewan into four quadrants [Southeast (SE), Southwest (SW), Northwest (NW), and Northeast (NE)] using a multistage sampling approach. This was followed by selection of rural municipalities (RMs) and small towns from these quadrants, employing Statistics Canada guidelines (89).

Details pertaining to the description of sampling approaches can be found elsewhere (87). Stage two comprised of administrating a questionnaire to the target population. In stage three, the selection of a sub-group from the target population was done to conduct clinical assessments. The

component of clinical assessments was not relevant to the current project. Individuals who participated in the baseline survey were followed for a period of four years.

For both time points (i.e. baseline survey, and follow-up survey), data on contextual as well as individual factors were gathered. My thesis is based on the baseline and follow-up data. We identified and separated all persons aged 18 or older into two groups, namely self-reported doctor diagnosed HRC cases and non-cases. The age cut-off point of 18 or older was utilized because the SRHS data set contains information on individuals aged 18 or older and was used as a secondary data set in this thesis.

Ethics approval was obtained for the SRHS from the Biomedical Research Ethics Board of University of Saskatchewan, Canada (Bio#09-56; approved on 9 April 2009). Ethics for the current thesis was received from the Biomedical Research Ethics Board of University of Saskatchewan, Canada (Bio#1789; approved on 6 March 2020) (Refer to Appendix-C).

3.1.2 Study Population and Data Collection

The Study population was made up of rural municipalities (RMs) and small-town farm and non-farm rural dwellers residing in four quadrants [Southeast (SE), Southwest (SW), Northeast (NE) and, Northwest (NW)] of Saskatchewan, Canada. Using a multistage sampling design, the SRHS sampled tax-paying households from RMs and towns of Saskatchewan (87). The southern half of the province of Saskatchewan was divided into quadrants using Statistics Canada guidelines (89).

Out of each of the four quadrants, 12 RMs were chosen. The number of towns and RMs of rural Saskatchewan that participated in the SRHS was; 16 of 145 and 48 of 297, respectively. Thirty-two 32 of 36 RMs [SE (7), SW (8), NE (8), and NW (9)], representing (89%) and, 15 of 16 small-towns [SE (3), SW (4), NE (2), and NW (6)] representing (94%) participated in the baseline survey (87). The baseline questionnaire was mailed to 11004 households (87), and of this number, data was collected on 8261 persons aged at least 18 years who were nested in 4624 households (90). Using a similar questionnaire in the follow-up study, data was collected on 4867 individuals aged 18 years or older who were also nested in 2797 households (90). Worthy of note is that 4741 individuals participated in both surveys while 126 participants who primarily did not originally participate in the baseline survey, now participated in the follow-up study.

For data collection, a household survey questionnaire was created and employed to gather data. The SRHS research team along with two members from the community formed the questionnaire. One of these community members was from a RM and the other member was from a town.

The response rates for both the baseline and follow-up surveys were maximized using a modified version of Dillman's method for mail and telephone surveys (91). Dillman's method involves monitoring and maintaining a series of mail correspondence with the prospective participants of the study.

Data about the individual (e.g. smoking and alcohol consumption), contextual factors (e.g. household smoking), and occupational exposures (e.g. radiation and pesticides), and important covariates (e.g. age and gender) were collected.

A copy of the baseline questionnaire can be found in the appendix section of the thesis (refer to appendix A). This thesis project was based on HRCs cases identified both at the baseline as well as the follow-up for determining the prevalence, incidence, and associated risk factors for the prevalence of HRCs in rural Saskatchewan.

3.1.3 Sample size

3.1.3.1 HRCs (Hormone-related cancers)

At baseline, our study sample consists of 8261 participants aged 18 or older who were at risk of developing HRCs. Of this number, 247 participants self-reported doctor diagnosed HRCs (i.e. 117 prostate and 130 breast cancer cases) while the remaining 8013 did not (i.e. non-cases). However, at the follow-up, the total number of participants was 4867, with HRCs cases of 167.

Of note is, the HRC cases (breast and prostate cancer combined) of 167 at the follow-up were inclusive of HRCs cases at the baseline.

3.1.3.2 Breast cancer

At baseline the sample size was consisted of 4195 females. Of this number, there were 130 cases of breast cancer and the remaining 4065 non-cases. Similarly, at follow-up, the sample size was made up of 2502 females. Of this number, there were 94 cases and 2408 non-cases.

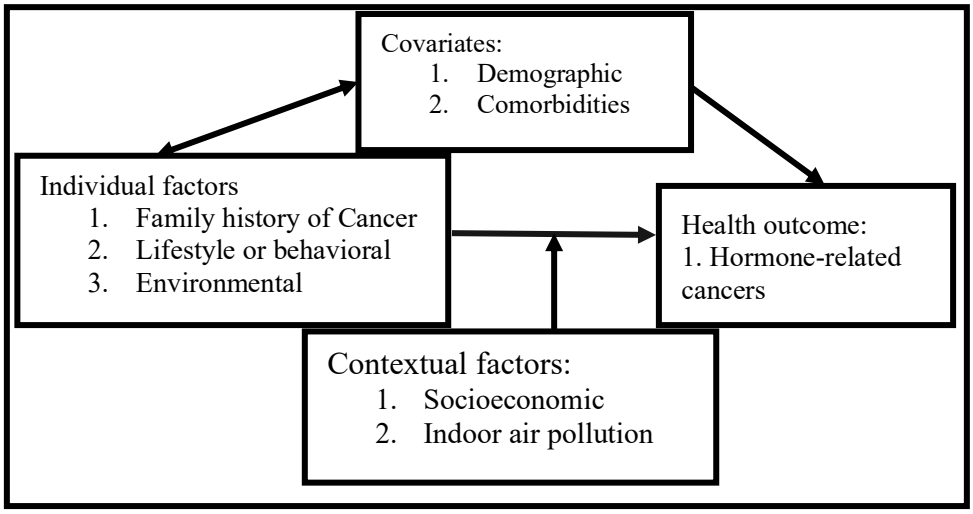
3.1.3.3 Prostate cancer

For prostate cancers, the sample size was composed of 4064 males at baseline. This number was made up of 117 cases of prostate cancer and 3947 non-cases. Similarly, the sample size for this cancer at follow-up was 2364 males, representing 73 cases and 2291 non-cases.

3.1.4 Theoretical framework

I will use the concept of Health Canada's "Population Health Framework" (PHF) (66) which was successfully implemented elsewhere (66) to help identify risk factors for the prevalence of HRCs in Saskatchewan. The PHF suggests that the interaction of individual and contextual factors may result in a varying risk of unfavorable health outcomes. Figure 1 is a graphical representation of the conceptual framework employed to establish how some risk factors (as may be contained in the SRHS data file) influence the risk of HRCs.

Figure 3.1: Conceptual framework (Adopted and modified from (66), courtesy of Dr. Will Pickett of Queen's University)



3.1.5 Variables description

The following section provides description of the variables used for the purposes of this thesis. Section 3.1.5.1 provides operational definitions on individual factors under subsection 3.1.5.1.1 which is followed by details on contextual factors in section 3.1.5.1.2 and description of principal covariates in section 3.1.5.1.3. The last section 3.1.5.1.4 discusses the outcome variable for this research project.

For this thesis, the variables examined were categorized into individual/lifestyle factors, contextual factors, and covariates. These variables were obtained from the baseline questionnaire with reclassifying a few of those, where required.

The individual/lifestyle factors (i.e. personal history of smoking, alcohol consumption, and education, etc.) referred to factors measuring individual exposures.

The contextual factors (i.e. household source of water, SES, etc.) were referring exclusively to the rural environment. The principal covariates used were (age, BMI, gender, and, marital status).

In view of this thesis, the outcome variable was, whether an individual had ever been diagnosed with HRCs by a primary caregiver or a doctor, was established by the baseline questionnaire.

3.1.5.1 Operational definitions

3.1.5.1.1 Individual factors

Personal history of smoking:

This variable collected information on the individual history of smoking using the baseline questionnaire. Based on question B-36, the following question was asked; *Have you ever smoked cigarettes?*" with responses "*Yes/No*". In the context of this thesis, the smoking variable was classified into three groups; never-smoker, ex-smoker, and current smoker. Smoking was a derived variable and was formed based on questions B-37 through B-45 (See Appendix A).

Alcohol consumption:

Information on the history and frequency of alcohol consumption was gathered using the baseline questionnaire. To obtain this, the following question was asked; "*During the past 12 months, how often did you drink alcoholic beverages?*". To carry out analysis for this thesis, a new 'alcohol' variable was created. To answer this question, the newly formed variable classified responses into 3 classes; "*Everyday"*, "*Never"*, and "*Occasionally*" (all the remaining classes were combined to form this class). This was obtained using question B-46 (See Appendix A).

Physical activity:

This variable reflected the personal history of physical activity that was obtained using the baseline questionnaire. The following question was asked; "*Do you exercise"?* with a response option as *"Yes/No".* This question was in reference to question B-27 on the baseline questionnaire (See Appendix A).

Education:

This established the educational status which was self-reported by the participants of the study. The question asked was; *What is your highest level of education?*" with response options "*< high school", "Completed high school", "Completed university",* and *"Completed post-secondary education other than above*". In the context of the current thesis, a new 'education' variable was

created. The newly formed variable had two classes "*Grade 12 or below*" and "*Beyond Grade 12*". For grade 12 or below, the "*< high school*" and "*Completed high school*" were combined. Beyond grade 12, was composed of the combination of "*Completed University*" and "*Completed postsecondary education other than above*". These were in accordance with question B-4 (See Appendix A).

Early-life exposure to farm:

This variable collected information regarding early-life exposure to a farming environment. To determine the respondent's exposure to farm, the following question was asked; *Have you ever lived on a farm?*" with response options "*Yes/No*". Besides, to establish respondent's exposure to farm in their first year of life; the following question was asked; "Did you live on a farm during your first year of life" with response options "Yes/No". These questions were in reference to questions; B-31 and B-32 on the baseline questionnaire (See Appendix A).

Personal and family history of cancer:

Utilizing the baseline questionnaire, the personal and family history of cancer was collected from question B-50, and B-56, respectively (See Appendix A). For personal history, the following question was asked; "*Has a doctor or primary caregiver ever said you have cancer?*" with response options "*Yes/No/Don't know*". For family history, participants were asked; "*Have the following members of your biological family ever had cancer?*" and the list included "*Father*, *Mother* and *Brother/Sister" with response options (Yes/No/Don't know)".*

Occupational rural exposure:

This variable collected information about respondent's exposure to chosen occupational exposures (i.e. livestock, stubble smoke, fungicides, herbicides, molds, oil/well fumes, radiation, solvent fumes, welding fumes, grain dust, wood dust, diesel fumes, asbestos dust, insecticides, mine dust, and other specify).

The question asked was; "*Have you been exposed to any of the following in your workplace?*" This question was based on B-58 (See Appendix A), from the baseline questionnaire. Also, exposure to Propane has been already dealt with in household fuel sources under contextual factors.

At this point, it is imperative to note that question (B-58) did not collect information on the specific types of exposure. This can be explained by considering the example of radiation exposure. While collecting information on this particular exposure (i.e. radiation), data were not

collected on whether it was exposure to solar radiation or machinery-related radiations, etc. However, all information collected on exposures was on work-related exposures (occupational exposures). If the response to the chosen work-related exposure was "Yes", the same question (i.e. B-58) gathered further information on the frequency of that exposure *(i.e. "If Yes, how often"?)*. This sub-question on frequency was further divided into daily, weekly, monthly and, occasionally and was coupled with another question that collected information on the duration of that particular exposure by asking for the number of years *(i.e. "how many years?").*

3.1.5.1.2 Contextual factors

Socioeconomic status (SES): For this variable, a proxy measure of household income adequacy was used. This variable was primarily derived from four groups that were formed following the definition of Statistics Canada (92). The total household income and the total number of people in the household were used as the basis for questions A-2 and A-20 (see Appendix A). Details on the four household income adequacy groups are presented below in a tabulated form.

Household Income Adequacy groups and relevant details

Number of groups	Size of household	Income Band	Income Adequacy
Group 1	1 or 2 persons	< $15,000	
	3 or 4 persons	< $20,000	Lowest income
	5 or more persons	< $30,000	
Group 2	1 or 2 persons	$15,000-$29,000	Lower-middle income
	3 or 4 persons	$20,000-$39,000	
	5 or more persons	$30,000-$59,000	
Group 3	1 or 2 persons	$30,000-$59,000	Upper-middle income
	3 or 4 persons	$40,000-$79,000	
	5 or more persons	$60,000-$79,000	
Group 4	1 or 2 persons	≥$60,000	Highest income
	3 or 4 persons	≥$80,000	

Household smoking:

This variable had two categories "Yes/No" which denoted whether any household member used tobacco-related products which included cigars, cigarettes, and, pipes. For the household smoking variable, the following question was asked: *"Do any of the people in your house use any of the*

following tobacco products in your home?" This was obtained from question A-17 using the baseline questionnaire. (See Appendix A).

Household fuel source:

Municipality referred to whether a household of interest was located in a small town or a rural municipality (RM) of Saskatchewan.

Household fuel source – Natural gas, and Propane: For collecting information on household fuel source, the baseline questionnaire was employed based on question A-7 (See Appendix A); the following question was asked; "*What are the types of fuel sources used to heat your home" "natural gas (Yes/No)" and "propane (Yes/No)".*

Mildew or musty smell: Using a questionnaire based on question A-14 (See Appendix A), information on the household indoor environment was gathered. The following question was asked to obtain information on variable mildew odor or musty smell; *Does your home (including basement) frequently have mildew or musty smell?*" with answers "*Yes/No*".

Quadrant: This was a categorical variable and represented the geographical location of the participant's households. This variable had four levels; SE, SW, NW, and, NE and were coded 1-4, respectively. Using the 2006 census subdivision, information on quadrants were collected. This was found in the baseline questionnaire.

Farm: The "Farm" variable indicated the location of the home and/or workplaces that were involved in the production of agricultural-related products for sale. To derive this variable, the following question was asked; "*Where is your home located*?" and with responses "*Farm, In-town, Acreage*". In the context of this thesis, the farm variable was categorized into "Non-farm" and "Farm". The dichotomization of the farm variable (i.e. non-farm and farm) was required as rural non-farming exposures are distinct as compared to rural farming exposures. The Non-farm variable was created by combining "In-town" and "acreage". This was obtained using the baseline questionnaire; question A-1 (See Appendix A).

Household source of water: Household water source had four classes. From the baseline questionnaire, referring to question A-37 (See Appendix A); it was asked that; *What is the source of water supply for drinking purposes in your home?*" with options of "*Bottled water", "Deep well water (more than 100 feet)", "Shallow well water (less than 100 feet)", Spring, river or creek, Dugout or reservoir, Lake, and Other source*". A new variable "Water source" was created which had the following categories "*Bottled water, deep well water (more than 100ft),*

and shallow well water (less than 100 ft.)". Also, all the remaining classes were combined with "Other source".

3.1.5.1.3 Covariates

Age: Data on individuals age 18 years or older were collected in the SRHS at baseline. For the purposes of this thesis, the same age cut-off point of 18 years or older will be used. We are using this age cut-off point because the SRHS data set contains information on individuals aged 18 or older and is being used as a secondary data set in this thesis. The age variable was continuous, and the baseline questionnaire was employed to collect information on this variable. Age was sub-categorized into 18-45 (reference category), 46-55, 56-65, and 65 or older.

BMI: To derive this variable, the baseline questionnaire was used. BMI was computed using the following formula: [Weight in kilograms/ (Height in centimeters)^2] x 100,000. Participants were questioned as "*What is your height? _kg of ft and in*" and "*What is your weight? _kg or lbs*". Respondents self-reported their weight and height. For carrying out analysis, the baseline BMI was obtained in three categories including; normal weight (reference category), (0 - <25kg/m2), overweight (25-30 kg/m2), and, obese (>30 kg/m2). These were based on B-6 and B-7 questions of the baseline questionnaire (see Appendix A).

Gender: This was a variable with two categories "Male" (reference category) and "Female". The question was based on B-3 (see Appendix A).

Marital status: To establish marital status, respondents were asked this question; "*What is your marital status?*". Respondents self-reported their status by answering this question. This question was based on question B-8 (see Appendix A).

3.1.5.1.4 Outcome variable

The primary outcome variable was self-reported, primary caregiver/doctor diagnosed, breast and prostate cancer (HRCs). The outcome of interest was achieved using the question from the baseline questionnaire. The question used in determining the outcome variable was *"Has a doctor or primary caregiver ever said you have cancer? If yes, please specify cancer type".*

For carrying out analysis, the "breast and prostate cancer" variable was recoded into a dichotomous variable *("Yes/No")*, such that breast and prostate cancer cases belong to the category "Yes", while breast and prostate cancer non-cases belong to the category "No". This was obtained in accordance with question B-50 (see Appendix A).

3.2 Statistical Methods

In this section, I present the methods used to analyze the data described in section 3.1 of this thesis. The SRHS data, as used herein was a longitudinal dataset. To achieve the objectives of this study, appropriate methods for longitudinal data analysis (LDA) must account for the within-cluster correlations that exist due to repeated measurements. In particular, marginal models were used to identify the risk factors for the prevalence of HRCs.

3.2.1 Analyzing Longitudinal Data

Very often in epidemiological studies, the aim is always to determine the nature of the relationship between an outcome and a set of independent variables. In cross-sectional studies, classical linear models or generalized linear models (GLMs) are used to determine such relationships (93). However, such models cannot be used for LDA because they do not account for the within-cluster correlations in longitudinal data since the independence assumption is violated (94). In this situation, appropriate models are developed to account for the inter-dependence or within-cluster correlations as a result of repeated measurements. One commonly used model for LDA is based on the generalized estimating equations (GEE) methodology (94). This methodology was applied in this thesis.

3.2.1.1 Statistical Application: Prevalence Analysis

3.2.1.1.1 Determining the Crude Prevalence

For this thesis, the prevalence may be defined as the "proportion of people in a population that has a disease" (96). Mathematically, this can be expressed as

$$\frac{Number\ of\ diseased\ people\ in\ the\ population}{total\ Population\ size} \text{ x } 100 \quad \text{.....(Eq 3.1)}$$

This formula was used to estimate the crude prevalence rates of breast and prostate cancers among the four quadrants at baseline and follow-up.

3.2.1.1.2 Determining Adjusted Prevalence

We determined the adjusted prevalence of breast and prostate cancers by fitting separate binary logistic regression models for covariates. Logistic regression models based on the Generalized Estimating Equations (GEE) methodology were fitted using SPSS to determine significant risk factors for the prevalence of breast and prostate cancers at baseline and follow-up respectively.

Generalized estimating equations (GEEs) approach is an iterative procedure (121), which was originally developed to analyse repeated measures data obtained from classical longitudinal studies. GEE is also used to analyze hierarchical data (i.e. units of analysis generally individuals (1^{st} level) are nested within a contextual level, for example household (2^{nd} level) (121). The repeated observations within one subject or individuals within household are not independent of each other, it is because of this reason, a correction should be made to account for within-subject correlations. This correction is possible by using GEE. GEE uses quasi-likelihood to estimate the regression coefficients (121).

GEE conducts correction by assuming a priori a certain 'working" correlation structure for the repeated measurements of the outcome variable (Y). The most commonly used working correlation structures are; a) Independent structure, b) Exchangeable structure, c) Stationary (m-dependent) structure, d) Auto-regressive correlation structure, and e) Unstructured correlation structure.

We used software SPSS and selected the exchangeable correlation matrix to account for hierarchy in the data. In exchangeable correlation structure only one correlation coefficient is estimated. GEE will account for clustering effect meaning by it will account for the several people with same characteristics within the cluster i.e. several people within same household (same people with same family might share the exposure).

For the purposes of this thesis, I used the statistical software IBM SPSS (version 26.0) to analyze data. As an example, and for the purposes of illustration, I have provided in the appendix a sample of SPSS syntax based on GEE approach that I used to build a binary regression model using HRCs (yes or no) as the response variable and diesel fumes (yes or no) as the predictor variable. (See Appendix- B)

A univariate and multivariate logistic regression model was used to investigate the association between breast and prostate cancer prevalence and a set of predictor variables. Multilevel (1^{st} level - individuals nested within 2^{nd} level - households) logistic regression models were used to assess the association between environmental/contextual and individual factors as well as covariates. Standard model building strategies were used to select the final multivariate model (97).

During the univariate analysis, variables with p-value<0.25 were selected for the multivariable modeling stage. Statistical significance was determined at 5%. Significant variables, as well as other biologically relevant variables, were retained in the final model even if the latter was not

significant. Interactions and confounding were checked for in the model building process. ORs and their associated 95% CIs were used to describe the strength of associations between outcome and predictor variables.

3.2.2 Statistical Application: Incidence Analysis

3.2.2.1 Determining the Crude Incidence

For the purpose of this thesis, we adopted the definition of cumulative incidence also called incidence proportion as "the number of new events of a specific disease during a specified period of time in a specified population" (100). Mathematically, this can be expressed as:

$$Cumulative\ Incidence = \frac{\text{Number of new cases in a specified period}}{\text{Number of people at risk in the specified population}} \quad \text{.....(Eq 3.2)}$$

CHAPTER 4 - RESULTS

4.1. Introduction

The main goal of this thesis was to determine the prevalence and crude incidence of self-reported doctor diagnosed HRCs (i.e. breast and prostate cancers) and the risk factors associated with the prevalence of these malignancies. While the methods used to analyze the data contained in this thesis were described in the previous section, this section presents the results of the data analysis.

4.2 Sample population

The SRHS data set contained information of 8,261 participants aged 18 or older who were nested within 4,624 households at baseline and 4,867 participants within 2,797 households during follow-up. Of these, 248 and 167 self-reported doctor diagnosed HRC cases were identified from the baseline and follow-up surveys respectively and included in during data analysis. For the purposes of this study, all these participants were included in the analysis because they were considered to be at risk of developing HRCs. It is worthy of note that, this study was particularly interested in prostate cancers among men and breast cancers among women. As a result, a single case of male breast cancer was identified and excluded from all analyses contained herein.

4.3. Descriptive Statistics

In this section, I provide descriptive statistics for the data used in this thesis. Table 4.1 presents the HRC cases and non-cases as stratified by the location of their homes (i.e. farm or non-farm residents). It is thus realized that 96 HRC cases, representing 38.9% were farm residents while the remaining 151 (61.1%) were non-farm residents. In total, only 3% of the study population reported HRC cancers while the remaining were treated as non-cases (Table 4.1). This 38:61 split in the proportion of HRC cancers stratified by farm and non-farm residents at baseline was similar to that of follow-up (Table 4.10).

Table 4.1 Location of Home by HRC cases and non-cases at Baseline

Location of home	HRC – Cases*	Non-Cases	Total
Farm	96 (38.9)	3349 (42.1)	3445 (42.0)
Non-Farm	151 (61.1)	4612 (57.9)	4763 (58.0)
Total	247 (3.0)	7961	8208

Tables 4.2 and 4.3 provide the population characteristics as stratified by the four quadrants at baseline and follow-up respectively. It thus appears that the four quadrants have a similar proportion of participants both at baseline and follow-up. At baseline, the proportion of female participants appears to be fairly higher than their male counterparts in all four quadrants, a relationship that persisted even at follow-up.

Study participants also appear not to be unevenly distributed across the various age groups considered in this study, especially within quadrants (Table 4.2). The proportion of older participants (> 65 years) was higher than all that for other age categories except in the north-western quadrant of Saskatchewan (Table 4.2). However, this distribution of study participants was wholly true at follow-up (Table 4.3).

Table 4.2 Population Characteristics for HRCs by Quadrant at Baseline

Characteristics[#]	Quadrant[*]			
	South West	**South East**	**North East**	**North West**
	N = 1538	N = 1792	N = 2400	N = 2527
Gender, n (%)				
Male	751 (48.8)	884 (49.3)	1198 (49.9)	1229 (48.7)
Female	787 (51.2)	908 (50.7)	1202 (50.1)	1295 (51.3)
Age-groups, n (%)				
18-45	343 (22.3)	465 (25.9)	486 (20.3)	650 (25.7)
46-55	432 (28.1)	441 (24.6)	553 (23.0)	622 (24.6)
56-65	320 (20.8)	421 (23.5)	599 (25.0)	606 (24.0)
>65	442 (28.8)	465 (25.9)	762 (31.8)	647 (25.6)
Location of home, n (%)				
Farm	552 (36.0)	704 (39.6)	1192 (50.0)	997 (39.7)
Non-Farm	980 (64.0)	1072 (60.4)	1193 (50.0)	1514 (60.3)

[#] Due to missing observations, variable totals may not sum to 8261

Table 4.3 Population Characteristics for HRCs by Quadrant at Follow-up

Characteristics[#]	Quadrant[*]			
	South West	**South East**	**North East**	**North West**
	N = 854	N = 1016	N = 1391	N = 1480
Gender, n (%)				
Male	414 (48.5)	486 (47.8)	681 (49.0)	724 (48.9)
Female	440 (51.5)	530 (52.2)	710 (51.0)	756 (51.1)
Age-groups, n (%)				
18-45	97 (11.4)	158 (15.6)	126 (9.1)	232 (15.7)
46-55	187 (21.9)	200 (19.7)	254 (18.3)	318 (21.5)
56-65	241 (28.2)	313 (30.8)	407 (29.3)	402 (27.2)
>=65	329 (38.5)	345 (34.0)	604 (43.4)	528 (35.7)
Location of home, *N* (%)				
Farm	319 (37.6)	415 (41.1)	700 (50.6)	612 (41.4)
Non-Farm	530 (62.4)	595 (58.9)	683 (49.4)	866 (58.6)

4.4. Prevalence Analysis

4.4.1 Crude Prevalence Analysis

As indicated earlier, one of the objectives of this thesis was to determine the prevalence of HRCs (i.e. breast and prostate cancers) both at baseline and follow-up. Data was analyzed to determine the crude prevalence of these cancers and the results are presented in the following tables (i.e. Tables 4.4 to 4.10).

Table 4.4 presents the prevalence of breast and prostate cancers combined among farm and non-farm residents and stratified by the various quadrants of Saskatchewan while Tables 4.5 and 4.6 contain that for prostate and breast cancers separately.

It is thus seen from Table 4.4 that, HRCs were slightly more prevalent amongst non-farm residents (3.2%) when compared with farm residents (2.8%). It is observed that HRCs were more prevalent within the eastern part of the province as compared to the western part (Table 4.4) among both farm and non-farm residents. Further research is needed to explore these observations. Analogous interpretations and observations were implied by Tables 4.5 and 4.6 for the prevalence of prostate and breast cancers, respectively for farm and non-farm residents.

It is refreshing to note that, the above commentary pertains to the baseline survey. For the avoidance of duplication in interpretations, Tables 4.7 through 4.9 provide similar information for HRCs (breast and prostate combined), prostate and breast (separately) respectively for the follow-up survey.

Table 4.4 Prevalence of HRCs Stratified by Geographic Location and Farm/Non-Farm Residence at Baseline.

	Farm Dwellers		Non-farm Dwellers		Total	
Quadrant of Saskatchewan	Doctor-diagnosed Self-reported HRCs		Doctor-diagnosed self-reported HRCs		Ever-diagnosed HRCs	
	Yes/Total	(%)	Yes/Total	(%)	Yes/Total	(%)
South West	15/552	2.7	35/980	3.6	50/1538	3.3
South East	22/704	3.1	33/1072	3.1	56/1792	3.1
North East	40/1192	3.4	42/1193	3.5	82/2400	3.4
North West	19/997	1.9	41/1514	2.7	60/2527	2.4
Total	96/3445	2.8	151/4763	3.2	248/8261	3.0

Table 4.5 Prevalence of Prostate Cancer Stratified by Geographic Location and Farm/Non-Farm Residence at Baseline.

	Farm Dwellers		Non-farm Dwellers		Total	
Quadrant of Saskatchewan	Doctor-diagnosed Self-Reported Prostate Cancer		Doctor-diagnosed Self-reported Prostate Cancer		Ever-diagnosed Prostate Cancer	
	Yes/Total	(%)	Yes/Total	(%)	Yes/Total	(%)
South West	8/296	2.7	16/452	3.5	24/751	3.2
South East	12/362	3.3	14/513	2.7	27/884	3.1
North East	22/626	3.5	20/565	3.5	42/1198	3.5
North West	10/508	2.0	14/712	2.0	24/1229	2.0
Total	52/1792	2.9	64/2244	2.9	117/4064	2.9

Table 4.6 Prevalence of Breast Cancer Stratified by Geographic Location and Farm/Non-Farm Residence at Baseline.

	Farm Dwellers		Non-farm Dwellers		Total	
Quadrant of Saskatchewan	Doctor-diagnosed Self-reported Breast Cancer		Doctor-diagnosed self-reported Breast Cancer		Ever-diagnosed Breast Cancer	
	Yes/Total	(%)	Yes/Total	(%)	Yes/Total	(%)
South West	7/256	2.7	18/528	3.4	25/787	3.2
South East	10/342	2.9	19/559	3.4	29/908	3.2
North East	18/566	3.2	22/628	3.5	40/1202	3.3
North West	9/489	1.8	27/800	3.4	36/1295	2.8
Total	44/1653	2.7	86/2517	3.4	130/4194	3.1

Table 4.7 Prevalence of HRCs Stratified by Geographic Location and Farm/Non-Farm Residence at Follow-up.

	Farm Dwellers		Non-farm Dwellers		Total	
Quadrant of Saskatchewan	Doctor-diagnosed self-reported HRCs		Doctor-diagnosed self-reported HRCs		Ever-diagnosed HRCs	
	Yes/Total	(%)	Yes/Total	(%)	Yes/Total	(%)
South West	9/319	14.5	22/530	21.6	31/854	18.9
South East	15/415	24.2	27/595	26.5	42/1016	25.6
North East	25/700	40.3	24/683	23.5	49/1391	29.9
North West	13/612	21.0	29/866	28.4	42/1480	25.6
Total	62/2046	3.0	102/2674	3.8	164/4741	3.5

Table 4.8 Prevalence of Prostate Cancer Stratified by Geographic Location and Farm/Non-Farm Residence at Follow-up.

	Farm Dwellers		Non-farm Dwellers		Total	
Quadrant of Saskatchewan	Doctor-diagnosed self-reported Prostate Cancer		Doctor-diagnosed self-reported Prostate Cancer		Ever-diagnosed Prostate Cancer	
	Yes/Total	(%)	Yes/Total	(%)	Yes/Total	(%)

South West	5/169	3.0	14/243	5.8	19/414	4.6
South East	6/214	2.8	11/269	4.1	17/486	3.5
North East	15/364	4.1	6/313	1.9	21/681	3.1
North West	6/314	1.9	10/408	2.5	16/724	2.2
Total	32/1061	3.0	41/1233	3.3	73/2305	3.2

Table 4.9 Prevalence of Breast Cancer Stratified by Geographic Location and Farm/Non-Farm Residence at Follow-up.

	Farm Dwellers		Non-farm Dwellers		Total	
Quadrant of Saskatchewan	Doctor-diagnosed self-reported Breast Cancer		Doctor-diagnosed self-reported Breast Cancer		Ever-diagnosed Breast Cancer	
	Yes/Total	(%)	Yes/Total	(%)	Yes/Total	(%)
South West	4/150	2.7	8/287	2.8	12/440	2.7
South East	9/201	4.5	16/326	4.9	25/530	4.7
North East	10/336	3.0	18/370	4.9	28/710	3.9
North West	7/298	2.3	19/458	5.1	26/756	3.4
Total	30/985	3.0	61/1441	4.2	91/2436	3.7

Table 4.10 Location of home by HRCs cases and non-cases at follow-up

Location of home	**HRC - Cases (%)**	**Non-Cases (%)**	**Total (%)**
Farm	64 (38.3)	2032 (43.4)	2096 (43.2)
Non-Farm	103 (61.7)	2646 (56.6)	2749 (56.7)
Total	167 (3.4)	4678 (96.6)	4845

4.4.1.1 Population Characteristics of HRCs

In the following section: potential risk factors for HRCs considered in this thesis are described. Table 4.11 provides the proportions of these characteristics at baseline and follow-up. In particular, the percentage of HRC cases and non-cases for the two time periods are provided in Table 4.11. Several contextual factors were included in the study. For socioeconomic status at baseline, a higher percentage of participants reported that they had "some money" (63.5% vs. 59.3%) left at the end of the month as compared to those who had “just enough money” and “not enough money”

amongst cases and non-cases respectively. Twenty-four (24) of 248 participants, representing 9.7% at baseline and 1,225 of 7968, representing 15.4% reported that there was household smoking in their homes while remaining 224 (90.3%) and 6743 (84.6) cases and non-cases (respectively) did not.

Also, almost 70% of HRC cases and non-cases reported natural gas as their main source of household fuel (Table 4.11). Only a small number of cases (16.4) and non-cases (17.1) had mildew or musty smell in their homes including the basement while the majority did not (83.6 vs. 82.9 respectively).

Amongst cases, about 35% of study participants reported bottled water as their household source of water while almost 15% reported lakes, springs, dugouts, etc. as their household source of water. While 27.2% of cases reported deep wells (more than 100 ft), 22.9% reported shallow wells (less than 100 ft) as their source of household water supply. A similar distribution was observed amongst non-cases (Table 4.11).

Three individual or lifestyle-related factors were considered in this study. These include; smoking status, alcohol consumption, and physical activity. The majority of cases (50.2%) and non-cases (52.7%) reported being "never smokers" while only 8.5% of cases and 11.9% of non-cases were current smokers at baseline. It is also important to note that a moderate percentage of cases (41.2%) and non-cases (35.4%) were ex-smokers.

A considerable number of cases (27.8%) reported never drinking alcohol while a few (8.5%) reported being daily consumers of alcohol. Similarly, amongst non-cases, only 10.6% were “everyday drinkers” of alcohol (Table 4.11). Regarding physical activity, a slightly higher percentage of both cases (59.2%) and non-cases (57.4%) self-reported engagement in any form of physical activity, while their counterparts did not (cases - 40.8%; non-cases - 42.6%).

Early childhood exposures were also explored in the present study. 90.7% of HRC cases and 82.6% of non-cases ever lived on a farm. In addition, while 79.1% of cases lived on a farm in their first year of life, 20.9% did not. A similar percentage distribution was observed amongst non-cases (Table 4.11).

A family history of cancer was also included in the current study. The cancer history of study participants' first-degree relatives (FDRs) was self-reported. All persons who reported breast or prostate cancers in this thesis also reported a previous history of malignancies other than HRCs (Table 4.11). A relatively lower percentage of cases reported that their fathers (38.1%), mothers

(32.6), and/or siblings (i.e. brother or sister) (45.3%) had a positive history of cancer. A similarly low percentage of participants who reported a positive history of cancer in either of their FDRs was observed amongst non-cases as well (Table 4.11).

Several occupational exposures were included in the current study. These include asbestos dust, diesel fumes, fungicides (to treat grain), grain dust, pesticides (to kill insects and weeds), livestock exposure, mine dust, molds, oil/gas well fumes, propane, radiation, stubble smoke, solvent and welding fumes. The distribution of the percentage of study participants, stratified cases and non-cases are provided in Table 4.11.

Besides, the current study explored some demographic factors including age, BMI, gender, level of education, and marital status. The study population both amongst cases and non-cases was predominantly older (>65 years). A slightly higher percentage of females (52.4%) than males (47.6%) cases recorded. However, amongst non-cases, it was a near 50-50 split (Table 4.11). The majority of both cases (42%) and non-cases (40.9%) were over-weighted. Data analysis also revealed that a higher percentage of both cases (71.9%) and non-cases (60.2%) completed grade 12 or lower. Many cases were also married or living together while a relatively small percentage of study participants were either divorced, widowed, separated or never married.

It is imperative to note that, the fore-running commentary on the statistics provided above describes the baseline survey and much be interpreted accordingly. Columns 3 and 4 of table 4.11 provide descriptive statistics for the follow-up survey and consequently, similar interpretations are implied.

Table 4.11 Population Characteristics for HRC Cases and Non-cases at Baseline and Follow-up

	Baseline (2010)		Follow-up (2014)	
Description	**Cases (N = 247) (%)**	**Non-cases (N = 8013) (%)**	**Cases (N = 167) (%)**	**Non-cases (N = 4700) (%)**
CONTEXTUAL FACTORS				
Socioeconomic				
Some money	139 (63.5)	4288 (59.3)	105 (68.2)	2853 (67.8)
Just enough money	39 (17.8)	1556 (21.5)	32 (20.8)	835 (19.8)
Not enough money	41 (18.7)	1387 (19.2)	17 (11.0)	522 (12.4)

Quadrant				
South West	50 (20.2)	1488 (18.6)	31 (18.9)	1507 (18.6)
South East	56 (22.6)	1736 (21.7)	42 (25.6)	1750 (21.6)
North East	82 (33.1)	2318 (28.9)	49 (29.9)	2351 (29.0)
North West	60 (24.2)	2467 (30.8)	42 (25.6)	2485 (30.7)
Household smoking				
Yes	24 (9.7)	1225 (15.4)	14 (8.4)	489 (10.4)
No	224 (90.3)	6743 (84.6)	153 (91.6)	4211 (89.6)
Fuel source – Natural gas				
Yes	172 (69.9)	5415 (67.8)	123 (74.5)	3137 (67.1)
No	74 (30.1)	2567 (32.2)	42 (25.5)	1540 (32.9)
Mildew or musty smell				
Yes	40 (16.4)	1328 (17.1)	17 (10.4)	747 (16.3)
No	204 (83.6)	6451 (82.9)	147 (89.6)	3839 (83.7)
Household source of water				
Bottled water	2415 (35.0)	71 (36.0)	46 (34.8)	1306 (34.1)
Deep well water (>100ft)	1874 (27.2)	62 (31.5)	41 (31.1)	1083 (28.2)
Shallow well water (<100ft)	1581 (22.9)	30 (15.2)	23 (17.4)	895 (23.3)
Others	1023 (14.8)	34 (17.3)	22 (16.7)	551 (14.4)
Location of home				
Farm	96 (38.9)	3349 (42.1)	64 (38.3)	2032 (43.4)
Non-farm	151 (61.1)	4612 (57.9)	103 (61.7)	2646 (56.6)
INDIVIDUAL FACTORS				
Smoking Status				
Current Smoker	21 (8.5)	947 (11.9)	12 (7.2)	430 (9.2)
Ex-smoker	102 (41.3)	2821 (35.4)	72 (43.4)	1773 (38.0)
Never smoker	124 (50.2)	4201 (52.7)	82 (49.4)	2461 (52.8)
Alcohol consumption				
Never	69 (27.8)	1403 (17.6)	44 (26.3)	790 (16.9)
Less than once a month	61 (24.6)	1679 (21.0)	32 (19.2)	955 (20.4)
At most 2-3 times a month	56 (22.6)	2136 (26.8)	39 (23.4)	1207 (25.8)
At most 2-3 times a week	41 (16.5)	1914 (24.0)	31 (18.6)	1119 (24.0)
Everyday	21 (8.5)	645 (10.6)	21 (12.6)	599 (12.8)
Physical activity				
Yes	138 (59.2)	4481 (57.4)	87 (52.7)	2494 (54.3)
No	95 (40.8)	3326 (42.6)	78 (47.3)	2103 (45.7)

Early life-exposures:				
Ever lived on a farm				
Yes	224 (90.7)	6598 (82.6)	149 (89.2)	4061 (86.5)
No	23 (9.3)	1392 (17.4)	18 (10.8)	634 (13.5)
Lived first year of life on farm				
Yes	193 (79.1)	5362 (67.6)	126 (76.4)	3370 (72.3)
No	51 (20.9)	2565 (32.4)	39 (23.6)	1293 (27.7)
Familial History of other cancers:				
Personal				
Yes	247 (100.0)	427 (5.4)	167 (100)	399 (8.6)
No	0 (0.0)	7478 (94.6)	0 (0.0)	4218 (91.4)
Father				
Yes	88 (38.1)	2038 (27.9)	69 (50.7)	1480 (39.3)
No	143 (61.9)	5254 (72.1)	67 (49.3)	2286 (60.7)
Mother				
Yes	73 (32.6)	1818 (24.4)	52 (40.3)	1356 (35.4)
No	151 (67.4)	5633 (75.6)	77 (59.7)	2470 (64.6)
Sibling(s)				
Yes	97 (45.3)	1518 (22.2)	76 (58.9)	1283 (36.6)
No	117 (54.7)	5321 (77.8)	53 (41.1)	2227 (63.4)
Occupational Exposures:				
At work, ever exposed to:				
Asbestos dust				
Yes	19 (7.9)	524 (6.7)	13 (8.0)	350 (7.7)
No	221 (92.1)	7332 (93.3)	149 (92.0)	4170 (92.3)
Diesel fumes				
Yes	142 (59.2)	4609 (58.7)	104 (63.8)	2906 (63.7)
No	98 (40.8)	3247 (41.3)	59 (36.2)	1658 (36.3)
Fungicides (to treat grain)				
Yes	99 (41.3)	2557 (32.5)	69 (43.1)	1808 (40.0)
No	141 (58.8)	5299 (67.5)	91 (56.9)	2710 (60.0)
Grain dust				
Yes	176 (73.3)	5347 (68.1)	129 (78.7)	3382 (73.7)
No	64 (26.7)	2509 (31.9)	35 (21.3)	1205 (26.3)
Pesticides				
Yes	85 (34.3)	3710 (46.3)	55 (32.9)	1954 (41.6)

No	163 (65.7)	4303 (53.7)	112 (67.1)	2746 (58.4)
Livestock				
Yes	137 (57.1)	4020 (51.2)	100 (61.7)	2727 (59.8)
No	103 (42.9)	3836 (48.8)	62 (38.3)	1837 (40.2)
Mine dust				
Yes	7 (2.9)	438 (5.6)	4 (2.5)	316 (7.0)
No	233 (97.1)	7418 (94.4)	158 (97.5)	4206 (93.0)
Molds				
Yes	82 (34.2)	2721 (34.6)	64 (40.0)	1869 (41.3)
No	158 (65.8)	5135 (65.4)	96 (60.0)	2657 (58.7)
Oil/Gas well fumes				
Yes	59 (24.6)	1885 (24.0)	51 (31.9)	1300 (28.8)
No	181 (75.4)	5971 (76.0)	109 (68.1)	3220 (71.2)
Propane use				
Yes	14 (5.7)	589 (7.4)	8 (4.8)	371 (7.9)
No	232 (94.3)	7393 (92.6)	157 (95.2)	4304 (92.1)
Radiation				
Yes	52 (21.7)	636 (8.1)	41 (25.5)	445 (9.9)
No	188 (78.3)	7220 (91.9)	120 (74.5)	4045 (90.1)
Stubble smoke				
Yes	109 (45.4)	3146 (40.0)	85 (52.1)	2083 (45.8)
No	131 (54.6)	4710 (60.0)	78 (47.9)	2465 (54.2)
Solvent fumes				
Yes	77 (32.1)	2788 (35.5)	57 (35.4)	1812 (40.1)
No	163 (67.9)	5068 (64.5)	104 (64.6)	2712 (59.9)
Welding fumes				
Yes	104 (43.3)	3223 (41.0)	71 (43.8)	2030 (44.8)
No	136 (56.7)	4633 (59.0)	91 (56.2)	2499 (55.2)
Wood dust				
Yes	90 (37.5)	3022 (38.5)	75 (46.3)	2051 (45.1)
No	150 (62.5)	4834 (61.5)	87 (53.7)	2498 (54.9)
Covariates				
Age (yrs.)				
18-45	3 (1.2)	1941 (24.2)	2 (1.2)	650 (13.8)
46-55	21 (8.5)	2027 (25.3)	8 (4.8)	982 (20.9)
56-65	57 (23.0)	1891 (23.6)	31 (18.6)	1355 (28.8)

>=65	167 (67.3)	2151 (26.9)	126 (75.4)	1713 (36.4)
BMI (kg/m^2)				
Normal (0-<25)	77 (32.4)	2268 (29.8)	50 (32.1)	1325 (29.8)
Overweight (25-30)	100 (42.0)	3107 (40.9)	63 (40.4)	1832 (41.2)
Obese (>30)	61 (25.6)	2228 (29.3)	43 (27.6)	1292 (29.0)
Sex				
Male	118 (47.6)	3946 (49.3)	73 (43.7)	2291 (48.8)
Female	130 (52.4)	4064 (50.7)	94 (56.3)	2408 (51.2)
Level of education				
<= Grade 12	174 (71.9)	4767 (60.2)	106 (64.6)	2636 (56.9)
> Grade 12	68 (28.1)	3150 (39.8)	58 (35.4)	2000 (43.1)
Marital status				
Married/common law/living together	185 (74.9)	6595 (82.7)	136 (81.4)	3983 (85.0)
Widowed/divorced/single/separated	62 (25.1)	1384 (17.3)	31 (18.6)	702 (15.0)

4.4.2 Adjusted Prevalence Analysis

4.4.2.1 Univariable analysis for HRCs

As described in Chapter 3 of this thesis, standard model building procedures were followed to select significant risk factors for the adjusted prevalence of HRCs. Tables 4.12, 4.14, and 4.16 present the univariable analysis of the association between HRCs (i.e. breast and prostate cancers combined), breast and prostate cancers (separately) and contextual, individual factors as well as covariates. Odds ratios are used to describe the associations and a p-value of 0.25 and 0.05 was used as the cut-off point for significance in the univariable and multivariable modeling respectively. The clustering effects of more than one individual within a household were accounted for using multilevel binary logistic regression based on generalized estimating equations (GEE) approach.

At baseline, significant predictors of HRCs included contextual (quadrant, household smoking, household water source), individual (smoking status and alcohol consumption), early life exposures (whether or not an individual ever lived on farm, living on farm during one's first year of life), family history of cancer (father, mother or sibling's positive history of cancer), occupational exposures (i.e. fungicides, grain dust, pesticides, exposure to livestock, mine dust, radiation, and stubble smoke), and other important covariates including age, educational and

marital status (Table 4.12). The last two columns of Table 4.12 respectively provide the unadjusted ORs and p-value for the univariable modeling of HRCs risk at follow-up. Details of the univariable analysis are presented below in Table 4.12.

TABLE 4.12 Univariable Analysis of the Relationship between HRCs (Prostate and Breast cancers combined) and Contextual Factors, Individual Factors, and Covariates by Odds Ratio (OR), 95% CI, and P-value by Time-point

Predictor	**Baseline (2010)**		**Follow-up (2014)**	
	Unadjusted		Unadjusted	
Description	Odds Ratio (OR)* (95% CI)	P-value#	Odds Ratio (OR)* (95% CI)	P-value#
CONTEXTUAL FACTORS				
Socioeconomic		0.36		0.87
Some money	1.00		1.00	
Just enough money	0.77 (0.54, 1.11)		1.04 (0.70, 1.55)	
Not enough money	0.91 (0.64, 1.23)		0.89 (0.52, 1.52)	
Quadrant		**0.16**		0.37
South West	1.00			
South East	0.96 (0.65, 1.42)		1.15 (0.72, 1.83)	
North East	1.05 (0.74, 1.50)		0.97 (0.62, 1.52)	
North West	0.72 (0.49, 1.06)		0.76 (0.49, 1.24)	
Environmental:				
Household smoking		**0.19**		0.40
Yes	1.46 (0.83, 2.55)		0.79 (0.46, 1.36)	
No	1.00		1.00	
Location of home		0.32		**0.19**
Farm	0.88 (0.67, 1.14)		0.81 (0.59, 1.11)	
Non-farm	1.00		1.00	
Water source		**0.01**		0.36
Bottled Water	1.00		1.00	
Deep well water	1.13 (0.79, 1.60)		1.08 (0.70, 1.65)	
Shallow well water	0.65 (0.42, 1.00)		0.73 (0.44, 1.21)	
Other sources	1.34 (0.99, 1.87)		1.12 (0.75, 1.68)	
Fuel source – Natural gas		0.49		**0.04**
Yes	1.10 (0.83, 1.45)		1.44 (1.01, 2.04)	

No	1.00		1.00	
Household Propane Use		0.35		**0.15**
Yes	0.76 (0.43, 1.35)		0.59 (0.29, 1.20)	
No	1.00		1.00	
Mildew odor or musty smell in home		0.77		**0.04**
Yes	0.95 (0.68, 1.33)		0.59 (0.36, 0.98)	
No	1.00		1.00	
INDIVIDUAL FACTORS				
Smoking Status		**0.09**		0.32
Current Smoker	0.75 (0.47, 1.20)		0.84 (0.46, 1.54)	
Ex-smoker	1.22 (0.94, 1.60)		1.22 (0.89, 1.68)	
Never smoker	1.00		1.00	
Alcohol consumption		**<0.001**		**0.03**
Never	1.00		1.00	
Less than once a month	0.74 (0.52, 1.06)		0.60 (0.38, 0.96)	
At most 2-3 times a month	0.54 (0.37, 0.77)		0.58 (0.37, 0.90)	
At most 2-3 times a week	0.44 (0.30, 0.65)		0.50 (0.31, 0.80)	
Everyday	0.51 (0.31, 0.83)		0.63 (0.37, 1.07)	
Physical activity		0.75		**0.70**
Yes	1.08 (0.83, 1.41)		0.94 (0.69, 1.28)	
No	1.00		1.00	
Early life-exposures:				
Ever lived on a farm		**<0.001**		
Yes	2.04 (1.33, 3.14)		1.29 (0.79, 2.11)	0.31
No	1.00		1.00	
Lived on a farm in first year of life		**<0.001**		
Yes	1.84 (1.33, 2.46)		1.24 (0.86, 1.78)	**0.25**
No	1.00		1.00	
Familial History of cancer:				
Father ever had cancer		**<0.001**		**0.001**
Yes	1.58 (1.21, 2.08)		1.59 (1.12, 2.25)	
No	1.00		1.00	
Mother ever had cancer		**<0.001**		0.26
Yes	1.49 (1.12, 1.98)		1.23 (0.86, 1.76)	
No	1.00		1.00	
Sibling(s) ever had cancer		**<0.001**		**<0.001**
Yes	2.90 (2.20, 3.83)		2.49 (1.75, 3.56)	

No	1.00		1.00	
Occupational Exposures:				
At work, ever exposed to:				
Asbestos dust	1.20 (0.73, 1.97)	0.47	1.04 (0.59, 1.85)	0.90
Diesel fumes	1.02 (.78, 1.32)	0.89	1.01 (0.73, 1.40)	0.97
Fungicides (to treat grain)	1.45 (1.12, 1.89)	**0.01**	1.14 (0.83, 1.57)	0.43
Grain dust	1.29 (0.97, 1.72)	**0.09**	1.31 (0.90, 1.91)	**0.16**
Pesticides (to kill plants and insects)	1.65 (1.28, 2.12)	**<0.001**	1.38 (0.99, 1.92)	**0.06**
Livestock	1.27 (0.98, 1.64)	**0.07**	1.09 (0.79, 1.50)	0.61
Mine dust	0.51 (0.24, 1.09)	**0.08**	0.34 (0.12, 0.91)	**0.03**
Molds	0.98 (0.75, 1.29)	0.89	0.95 (0.69, 1.31)	0.75
Oil/Gas well fumes	1.04 (0.77, 1.40)	0.82	1.12 (0.82, 1.63)	0.40
Radiation	3.14 (2.28, 4.33)	**<0.001**	3.11 (2.14, 4.53)	**<0.001**
Stubble smoke	1.24 (0.96, 1.61)	**0.10**	1.29 (0.94, 1.77)	**0.11**
Solvent fumes	0.86 (0.65, 1.13)	0.28	0.82 (0.59, 1.14)	**0.24**
Welding fumes	1.10 (0.85, 1.42)	0.47	0.96 (0.70, 1.32)	0.81
Wood dust	0.96 (0.74, 1.25)	0.76	1.05 (0.77, 1.44)	0.76
Covariates				
Age (yrs.)		**<0.001**		**<0.001**
18-45	1.00		1.00	
46-55	6.70 (2.0, 22.50)		2.49 (0.53, 11.77)	
56-65	19.50 (6.10, 62.37)		7.00 (1.67, 29.31)	
>65	50.23 (16.02, 157.49)		22.50 (5.56, 91.13)	
BMI (kg/m^2)		0.44		0.82
Normal (0-<25)	1.00		1.00	
Overweight (25-30)	0.95 (0.70, 1.28)		0.91 (0.62, 1.33)	
Obese (>30)	0.81 (0.56, 1.13)		0.88 (0.58, 1.34)	
Education		**<0.001**		**0.05**
≤ Grade 12	1.68 (1.27, 2.23)		1.39 (1.00, 1.92)	
> Grade 12	1.00		1.00	
Marital status		**<0.001**		**0.21**
Married/common law/living together	0.63 (0.47, 0.84)		0.77 (0.52, 1.15)	
Widowed/Divorced/separated/single/never married	1.00		1.00	

Significant risk factors at $p < 0.25$ in the univariable analysis were candidates for multivariable modeling. In tandem with established model building procedures, biologically relevant variables such as BMI, although not significant in the univariable modeling stage, were still included in the multivariable models. This is presented next.

4.4.2.2 Multivariable analysis for HRCs

After the multivariable adjustment, the significant predictors for the adjusted prevalence of HRCs at baseline were; the household source of water supply, father and sibling positive history of cancer, exposure to mine dust, and radiation, and age (Table 4.13-A). From the final multivariable model at baseline, the association between household source of water supply and HRCs risk was that of borderline significance. Our analysis revealed that individuals who self-reported their source of water supply as being deep well water were at lower risk of developing HRCs as compared to bottled water. A significant relationship was observed between the consumption of shallow well water and the HRCs risk. In particular, shallow well was associated with a reduced risk of HRCs. Our analysis further revealed that other sources of water (including lakes, dugouts, springs, etc.) were associated with a reduced HRCs risk although this association was not statistically significant and must be interpreted with caution. This latest finding was inconsistent with our hypothesis that lake/spring or dugout water would be associated with higher HRCs risk.

A family history of cancer was explored in multivariable modeling. We found that having a father with a positive history of cancer was associated with an elevated risk of being diagnosed with HRCs. This association was however recorded to be of borderline significance (p-value=0.05). Thus, for persons having a father with a previous history of cancer were 1.4 times more likely to be diagnosed with HRCs as compared to their counterparts with parents who had no such history (OR=1.37; 95% CI = 1.01, 1.86).

The data revealed a similar association between having a sibling (i.e. brother or sister) with a previous history of cancer and HRCs risk (Table 4.13-A). Thus, having a sibling with a positive history of cancer was associated with an almost doubled risk of developing HRCs as compared to having a sibling with no such history and this was statistically significant (OR = 1.51, 95% CI: 1.11, 2.07). Exposure to mine dust was associated with a significantly lower risk of developing HRCs (OR = 0.33; 95% CI: 0.12, 0.92) as compared to persons who were not. This finding might be due to a smaller number of observations.

It is important to note that, the current study revealed that, exposure to radiation was strongly and significantly associated with an increased risk of being diagnosed with HRCs. Thus, for individuals exposed to work-related radiations were associated with a more-than-three-times higher risk of developing HRCs as compared to individuals who were not exposed to radiation (OR = 3.39; 95% CI: 2.23, 4.84) (Table 4.13-A).

Covariates were also included in the multivariable adjustment. A significant dose-response relation was observed between age and HRCs risk. One must be very cautious when interpreting this finding/association, as confidence intervals appear to be very wide (Table 4.13-A). Although BMI, educational, and marital status qualified to be included in the final multivariable model as contained, they did not significantly predict the adjusted prevalence of HRCs at baseline (Table 4.13-A). The current study did not find any significant interactions among significant risk factors for the adjusted prevalence of HRCs and as a result, no interactions were reported.

The discussion provided above describes the situation for baseline analysis. The sixth column of Table 4.13-A provides the adjusted ORs associated with risk factors for the adjusted prevalence of HRC risk at follow-up and may be interpreted analogously. Moreover, the above discussion relative to risk factors for the adjusted prevalence of HRCs was done for the combination of breast and prostate cancers, called HRCs in this thesis. One must, therefore, exercise caution when interpreting or using these results as they pertain to HRCs (breast and prostate cancers combined).

TABLE 4.13-A: Multivariable Analysis of the Relationship between HRCs (Prostate and Breast cancers combined) and Contextual Factors, Individual Factors, and Covariates by Odds Ratio (OR), 95% CI, and P-value by Time-point

Predictor	**Baseline (2010)**			**Follow-up (2014)**		
Description	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Adjusted Odds Ratio (OR)* (95% CI)	P-value	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Adjusted Odds Ratio (OR)* (95% CI)	P-value
CONTEXTUAL FACTORS						
Location of home						0.67
Farm				-1.12 (0.27)	0.89 (0.52, 1.52)	

Non-farm		1.00		1.00	
Water source			**0.04**		
Bottled Water		1.00		1.00	
Deep well water	-0.12 (0.20)	0.88 (0.60, 1.31)			
Shallow well water	-0.71 (0.27)	0.59 (0.29, 0.82)			
Other sources	-0.10 (0.20)	0.99 (0.67, 1.46)			
Fuel source – Natural gas					0.27
Yes			0.36 (0.33)	1.43 (0.75, 2.71)	
No		1.00		1.00	
Household Propane Use					0.47
Yes			0.36 (0.49)	1.43 (0.54, 3.75)	
No		1.00		1.00	
Mildew odor or musty smell in home					0.24
Yes			-0.41 (0.36)	0.66 (0.33, 1.32)	
No		1.00		1.00	
INDIVIDUAL FACTORS					
Alcohol consumption			0.72		0.54
Never	Ref	1.00		1.00	
Less than once a month	-0.02 (0.22)	0.98 (0.64, 1.49)	-.50 (0.34)	0.61 (0.31, 1.19)	
At most 2-3 times a month	-0.18 (0.23)	0.83 (0.53, 1.31)	-0.39 (0.32)	0.67 (0.36, 1.27)	
At most 2-3 times a week	-0.19 (0.25)	0.83 (0.51, 1.34)	-0.16 (0.33)	0.85 (0.44, 1.64)	
Everyday	-0.37 (0.30)	0.70 (0.38, 1.26)	-0.05 (0.34)	0.95 (0.49, 1.185)	
Early life-exposures:					
Ever lived on a farm			0.74		
Yes	0.11 (0.34)	1.12 (0.58, 2.16)			

No		1.00				
Lived on a farm in first year of life			0.74			0.67
Yes	-0.08 (0.24)	0.92 (0.56, 1.48)		-0.12 (0.30)	0.88 (0.50, 1.58)	
No		1.00			1.00	
Familial History of cancer:						
Father ever had cancer			**0.05**			0.65
Yes	0.31 (0.16)	1.37 (1.01, 1.86)		0.10 (0.22)	1.10 (0.72, 1.68)	
No		1.00			1.00	
Mother ever had cancer						0.95
Yes				0.01 (0.23)	1.01 (0.65, 1.58)	
No					1.00	
Sibling(s) ever had cancer			**0.001**			0.30
Yes	0.42 (0.16)	1.51 (1.11, 2.07)		0.22 (0.21)	1.25 (0.82, 1.90)	
No		1.00			1.00	
Occupational Exposures:						
At work, ever exposed to:						
Grain dust	0.10 (0.22)	1.11 (0.71, 1.72)	0.65	0.10 (0.33)	1.10 (0.58, 2.01)	0.77
Pesticides (to kill plants and insects)	0.22 (0.18)	1.25 (0.87, 1.79)	0.23	-0.20 (0.30)	0.82 (0.45, 1.47)	0.49
Livestock	-0.14 (0.20)	0.87 (0.59, 1.28)	0.48			
Mine dust	-1.12 (0.53)	0.33 (0.12, 0.92)	**0.03**	-1.13 (0.73)	0.32 (0.08, 1.34)	0.12
Molds						
Oil/Gas well fumes						
Radiation	1.19 (1.20)	3.29 (2.23, 4.84)	**<0.001**	1.18 (0.27)	3.25 (1.90, 5.57)	**<0.001**
Stubble smoke	-0.18 (1.81)	0.84 (0.59, 1.19)	0.32	0.15 (0.27)	1.16 (0.68, 1.97)	0.59
Solvent fumes				-0.66 (0.27)	0.52 (0.31, 0.87)	**0.01**
Covariates						

Age (yrs.)			**<0.001**		**<0.001**
18-45	Ref.	1.00		Ref.	1.00
46-55	1.70 (0.63)	5.50 (1.61, 18.81)		0.53 (0.84)	1.70 (0.33, 8.84)
56-65	2.74 (0.60)	15.54 (4.78, 50.45)		1.45 (0.75)	4.26 (0.98, 18.51)
>=65	3.44 (0.59)	31.03 (9.68, 99.42)		2.71 (0.73)	15.09 (3.61, 63.11)
BMI (kg/m^2)			0.42		0.82
Normal (0-<25)		1.00			1.00
Overweight (25-30)	-0.10 (0.18)	0.90 (0.63, 1.29)		-0.09 (0.30)	0.99 (0.55, 1.78)
Obese (>30)	-0.12 (0.20)	0.89 (0.60, 1.32)		-0.20 (0.34)	0.82 (0.42, 1.60)
Education			0.17		0.27
≤ Grade 12	0.24 (1.79)	1.28 (0.90, 1.81)		0.26 (0.24)	1.30 (0.81, 2.09)
> Grade 12		1.00			1.00
Marital status			0.40		0.67
Married/common law/living together	-0.16 (1.94)	0.85 (0.58, 1.24)		0.13 (0.29)	1.14 (0.64, 2.01)
Widowed/Divorced/separated/single/never married		1.00			1.00

For the purposes of computing prediction probabilities, after the univariable analysis of HRCs, variables at $p<0.25$ and biologically or clinically relevant variables were selected. We fitted a model based on all the variables selected using the purposeful selection technique. Since some of the variable such as; age, quadrant, home location (farm, non-farm), were purposeful. Therefore, we retained all those variables in the model, even those were not significant at univariate or multivariate steps.

As a result, we obtained a full model containing twenty variables. Now, we selected variables at $p<0.05$ and biologically or clinically relevant variables to get a reduced model. The reduced model thus obtained comprised of six variables (i.e. age, quadrant, home location, sibling cancer,

radiation, and mine dust). Here, the variables like quadrant, and home location were found, not significant. However, history of sibling cancer, age, and work-related exposure to radiation remained significant.

Next, again at $p<0.05$ we selected variables to obtain the final model. Five variables qualified for this. In the final model the significant predictors for the adjusted prevalence of HRCs at the baseline were; age, quadrant, history of sibling cancer, and exposure to radiation. It was observed that age, exposure to radiations, history of sibling cancer, and quadrants independently contributed to increased risk of HRCs (Table 4.13-B).

At this point, it is refreshing to note that, same steps were followed to carry out univariate, and multivariate analysis of HRCs at the follow-up, and for breast & prostate cancer separately, at the baseline and follow-up, respectively.

For HRCs at the follow-up, our final model revealed that the significant predictors for the adjusted prevalence of HRCs were; age, quadrant, and work-related exposure to radiation (Table 4.13-B)

TABLE 4.13-B: Prediction Probabilities; Multivariable Analysis of the Relationship between HRCs (Prostate and Breast cancers combined) and Contextual Factors, Individual Factors, and Covariates by Odds Ratio (OR), 95% CI, and P-value by Time-point

Predictor	**Baseline (2010)**			**Follow-up (2014)**		
Description	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Adjusted Odds Ratio (OR)* (95% CI)	P-value	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Adjusted Odds Ratio (OR)* (95% CI)	P-value
CONTEXTUAL FACTORS						
Location of home						
Farm	.08(0.15)	1.08 (0.81, 1.45)	0.59	-0.10(0.16)	0.90(0.65, 1.25)	0.54
Sibling(s) ever had cancer						
Yes	0.42(0.15)	1.52 (1.13, 2.03)	0.005		----------	------
Occupational Exposures:						

At work, ever exposed to:						
Radiation	1.08(0.18)	2.94(2.06, 4.19)	**<0.001**	1.05(0.19)	2.87(1.96, 4.20)	**<0.001**
Covariates						
Age (yrs.)						
18-45	Ref.	1.00		Ref.	1.00	
46-55	1.83 (0.62)	6.23 (1.87, 21.03)	**0.003**	0.96(0.79)	2.6 (0.5, 12.4)	0.22
56-65	2.83 (0.59)	17.06 (5.33, 54.55)	**<0.001**	1.92(0.73)	6.8 (1.6, 28.6)	**0.009**
>=65	3.67 (0.58)	39.28 (12.41, 124)	**<0.001**	3.08(0.72)	21.7 (5.3, 88.5)	**<0.001**
Quadrant						
South West	0	1		0	1	
South East	-0.138 (0.22)	0.871 (0.56, 1.35)	0.540	0.32 (0.25)	1.4 (0.8, 2.3)	0.207
North East	-0.123 (0.19)	0.885 (0.60, 1.30)	0.530	-0.03 (0.24)	0.97 (0.6, 1.5)	0.888
North West	-0.425 (0.21)	0.654 (0.43, 0.99)	**0.05**	-0.04 (0.25)	0.003 (0.001, 0.01)	**<0.001**

***-----** indicates that the corresponding variables were not found significant at the final model although we retained the purposeful or otherwise biologically relevant variables in the model.

The probability of getting HRCs with age, and with work related exposure to radiations at the baseline are shown below in the figure 4.1a, and graph 4.1b, respectively.

Figure 4.1a: HRCs and Age (Baseline)

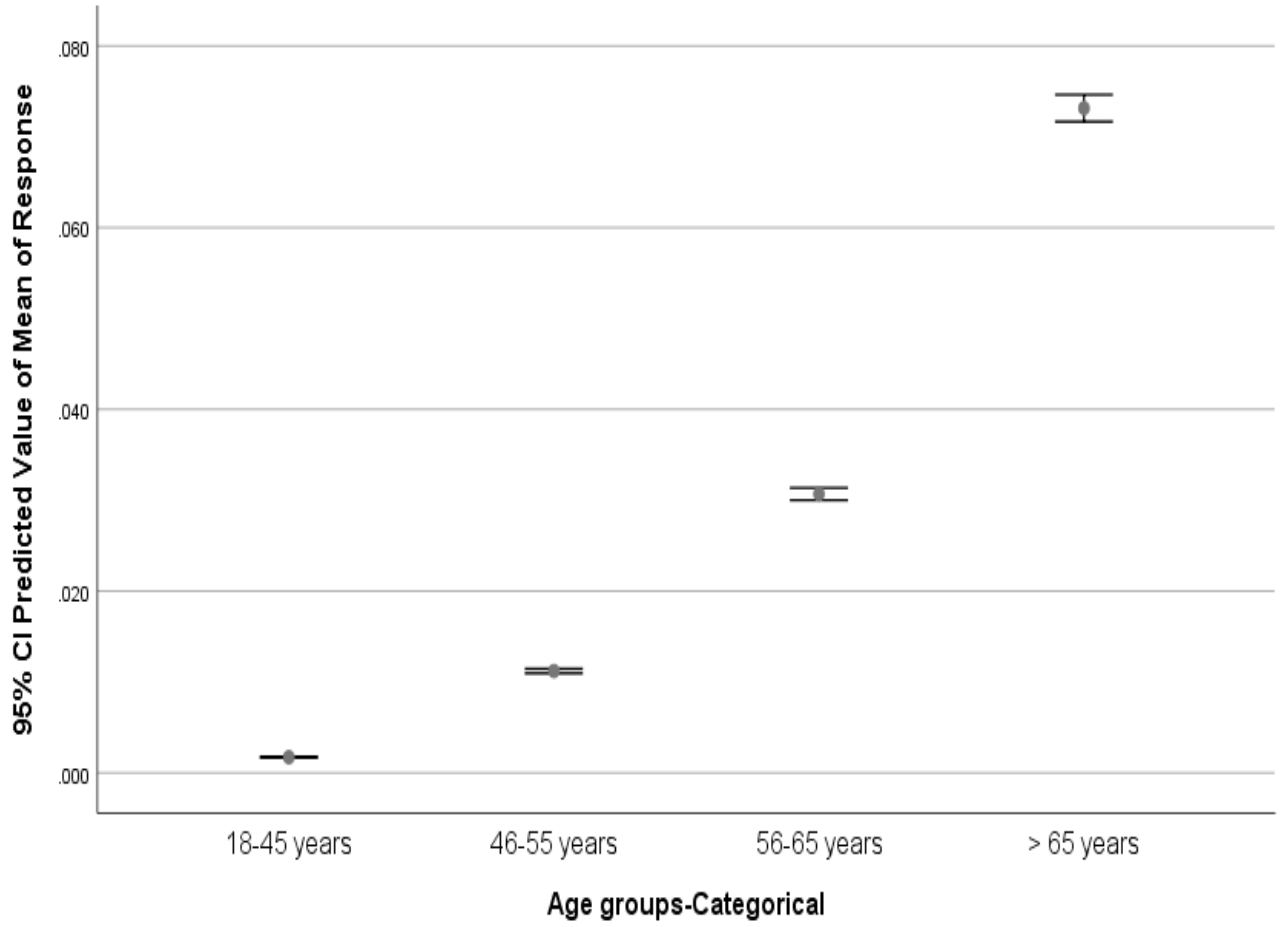


Figure 4.1a, indicates the probability of getting HRCs with age at the baseline.

Overall, there was a clear dose-response probability of having HRCs with increase in the age (i.e. 0.17% to 7.31%).

In terms of discrete age groups; the probability of having HRCs was 0.17% in the age group of 18-45 years, however for the age group 46-55 years, the probability constituted 1.12%, whereas for the age group 56-65 years this was 3.07%. Finally, for age group >65 years the probability reached to 7.31%. (Figure 4.1a)

Figure 4.1b: HRCs and Exposure to Radiation (Baseline)

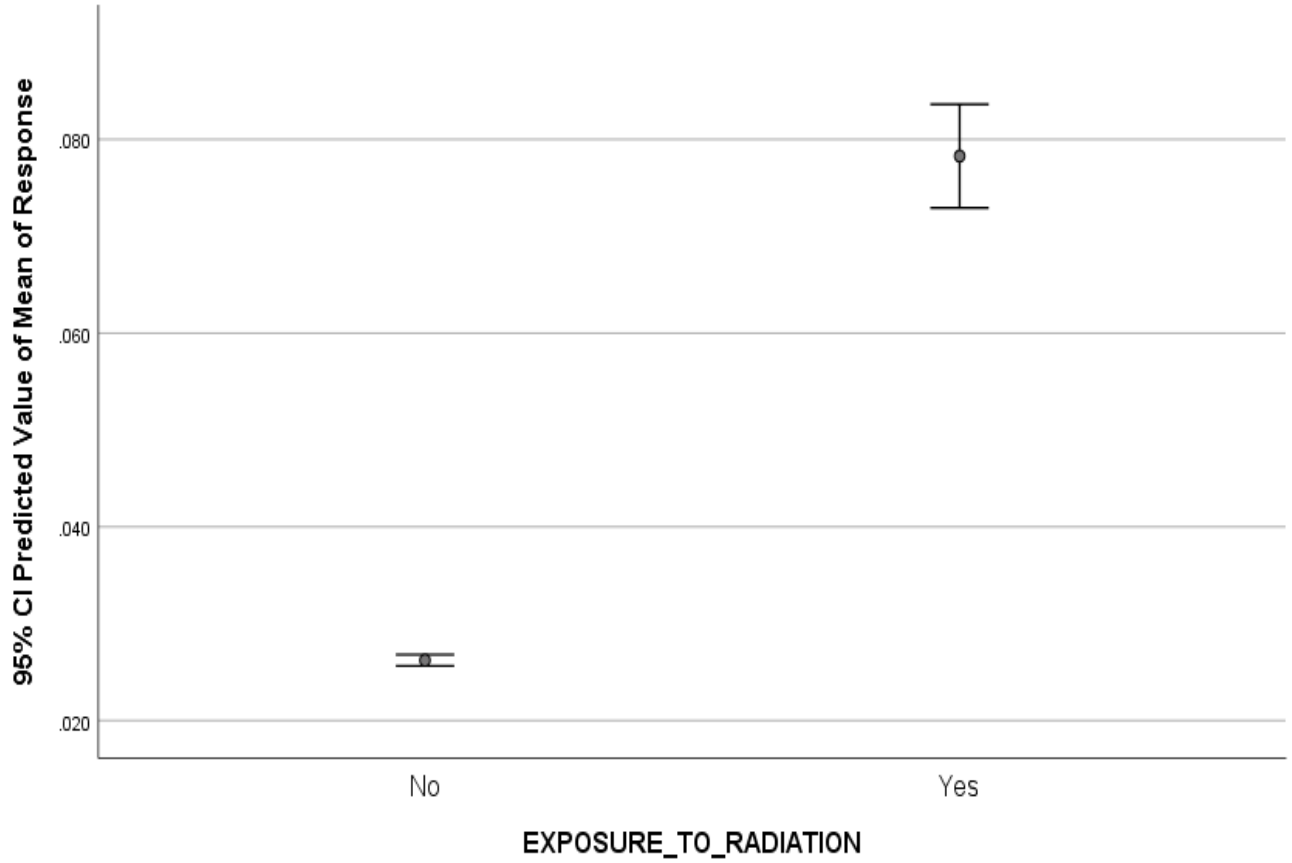
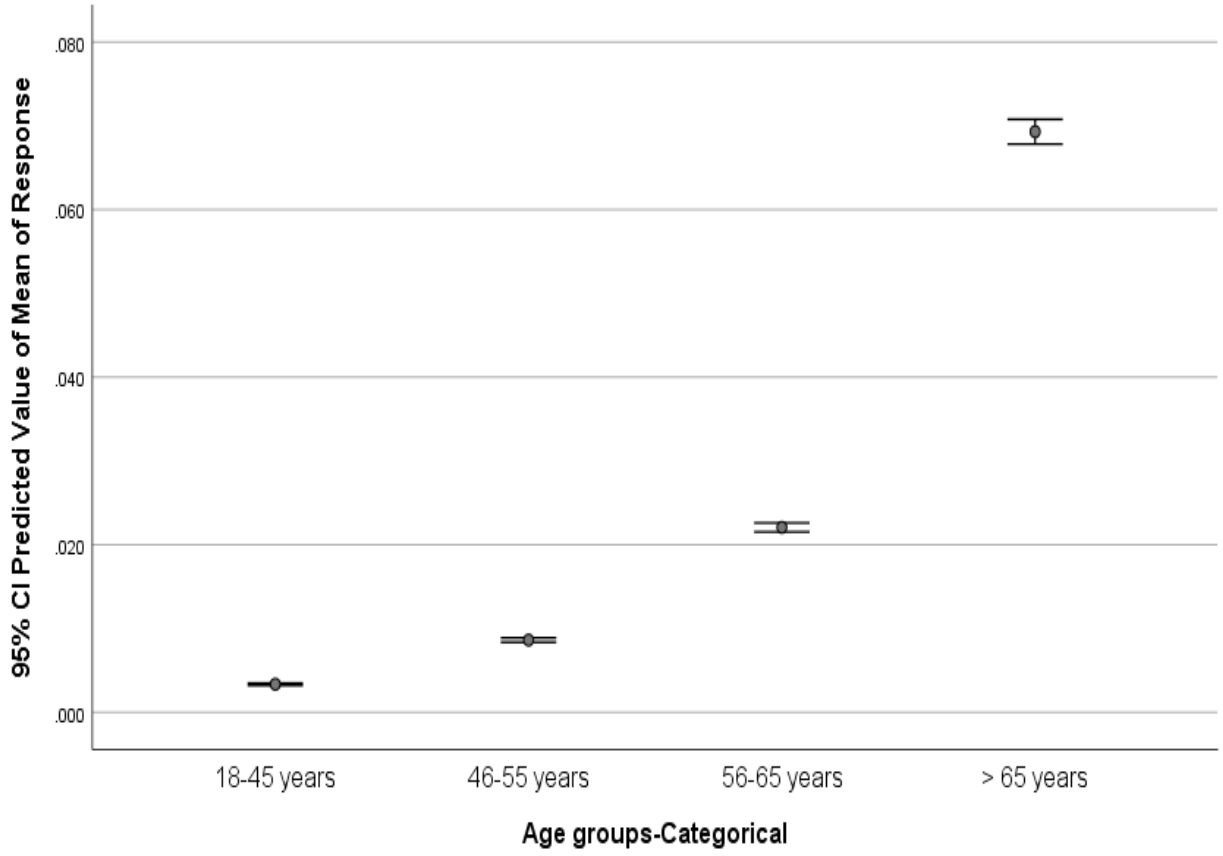


Figure 4.1b, shows the probability of getting HRCs with work related exposure to radiation at the baseline.

It is clear from the above graph that the probability of getting HRCs among those exposed to radiations was higher (7.83%) than those who were not exposed (2.62%).

Next, the probability of getting HRCs with age, and work-related exposure to radiations at the follow up is shown below in the figure 4.2a, and figure 4.2b, respectively.

Figure 4.2a: HRCs and Age (Follow-up)



The figure 4.2a above shows the probability of getting HRCs with age at the follow up.

Overall, it was observed that there was a clear dose-response probability of having HRCs with increase in the age (i.e. 0.50% to 7.03%).

In terms of discrete age groups, the probability of having HRCs was 0.50% in age group of 18-45 years, however for age group 46-55 years the probability constituted 1.5%, whereas for age group 56-65 years the probability was 3.95%. Finally, for the age group >65 years, this reached to 7.03% (Figure 4.2a).

Graph 4.2b: HRCs and Exposure to Radiations (Follow-up)

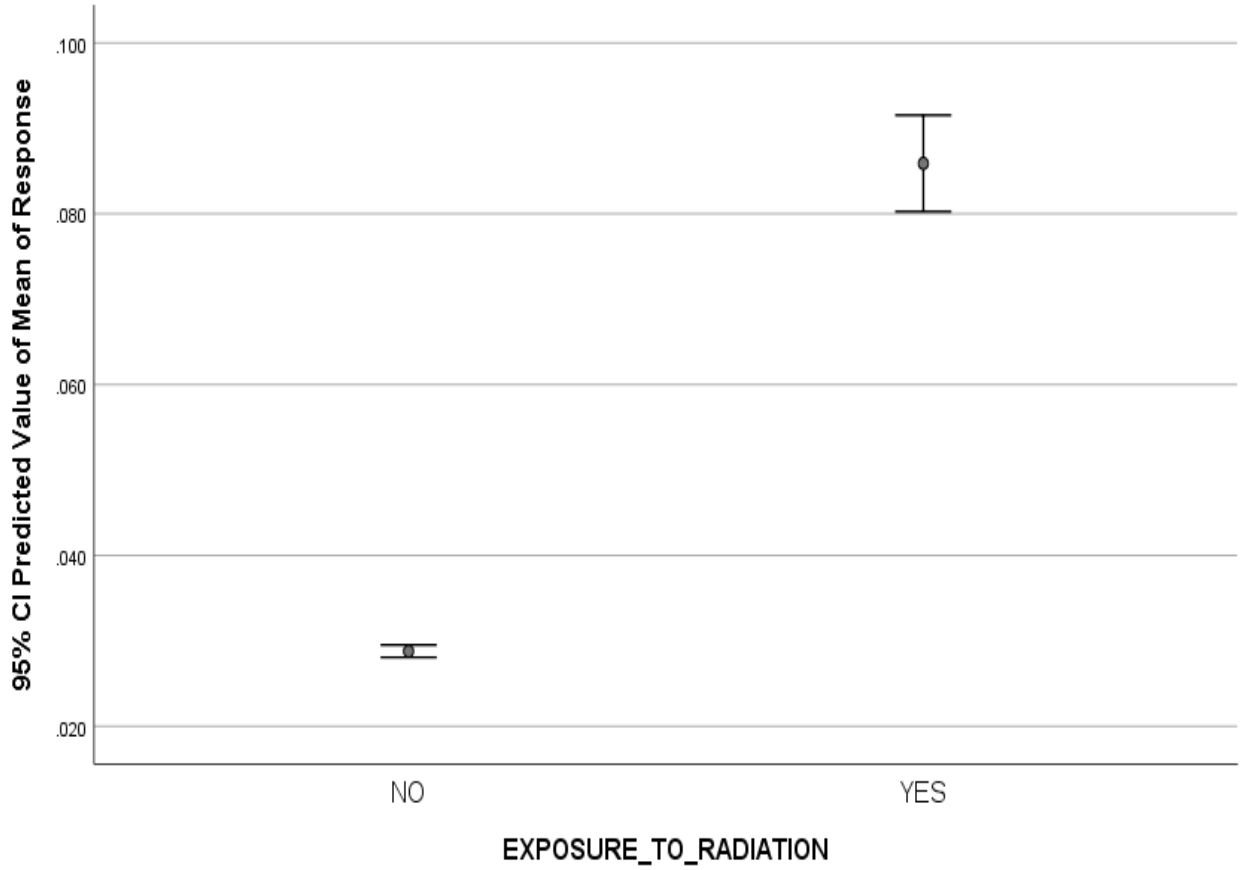


Figure 4.2b above shows the probability of getting HRCs with work related exposure to radiation at the follow up. It is evident from the figure 4.2b, that the probability of getting HRCs among those exposed to radiations were greater (8.6%) than those who were not exposed (2.9%), at the follow up.

It is imperative to note that, both figures (i.e. figure 4a, and figure 4b) for HRCs and exposure to radiations at the follow up expressed the same trend that was observed earlier for HRCs and age in (figure 4.1a, and, figure 4.1b), at the baseline. This implied that age continued to express a clear does-response relationship with having HRCs, meaning by as the age increased the probability of having HRCs increased. Similarly, for work related exposure to radiations, the same pattern was observed i.e. the probability of getting HRCs remained higher in those who

were exposed to work-place radiations as compared to those who were not, both at the baseline and at the follow up.

Next, tables 4.14 and 4.15-A respectively, describe the univariable and multivariable associations between potential risk factors for the adjusted prevalence of breast cancer. This will be followed by model building for prediction probabilities of the breast cancer at base line and follow up, results of those are shown in Table 4.15-B.

4.4.2.3 Univariable Analysis of Breast Cancer

The univariable associations between breast cancer and contextual factors, individual factors, and covariates were assessed. A p-value<0.25 was used as the cut-off to define significance. Table 4.14 presents the results of these univariable associations at both baseline and follow-up. Significant predictors of breast cancer at baseline include household smoking, location of home (farm or non-farm), household use of propane, mildew or musty smell in home, alcohol consumption, ever living on farm, living on farm in first year of life, family history of cancer (i.e. father, mother, and sibling's positive history of cancer), occupational exposure to asbestos, grain dust, pesticides, radiation and age. A significant dose-response relationship was noted between age and breast cancer risk (Table 4.14). At follow-up, significant predictors of breast cancer include; location of home (farm or non-farm), household source of water, household use of natural gas or fuel source, and propane use, mildew or musty smell in home, living on farm in first year of life, sibling's positive history of cancer, occupational exposure to grain dust, pesticides, radiation, age, educational and marital statuses. A significant dose-response relationship was also observed between age and the odds of developing breast cancer.

TABLE 4.14 Univariable Analysis of the Relationship between Breast Cancer and Contextual Factors, Individual Factors, and Covariates by Odds Ratio (OR), 95% CI, and P-value by Time-point

Predictor	**Baseline (2010)**		**Follow-up (2014)**	
Description	Unadjusted Odds Ratio (OR)* (95% CI)	P-value#	Unadjusted Odds Ratio (OR)* (95% CI)	P-value#

CONTEXTUAL FACTORS				
Socioeconomic		0.58		0.88
Some money	1.00		1.00	
Just enough money	0.78 (0.48, 1.26)		0.89 (0.51, 1.57)	
Not enough money	0.91 (0.56, 1.47)		1.08 (0.57, 2.04)	
Quadrant		0.87		0.41
South West	1.00		1.000	
South East	1.01 (0.58, 1.73)		1.77 (0.88, 3.55)	
North East	1.05 (0.63, 1.74)		1.46 (0.74, 2.91)	
North West	0.87 (0.52, 1.46)		1.27 (0.64, 2.54)	
Environmental:				
Household smoking		**0.19**		0.58
Yes	0.69 (0.39, 1.20)		0.81 (0.39, 1.69)	
No	1.00		1.00	
Location of home		**0.17**		0.18
Farm	0.77 (0.53, 1.12)		0.74 (0.48, 1.14)	
Non-farm	1.00		1.00	
Water source		0.28		0.10
Bottled Water	1.00		1.00	
Deep well water (more than 100ft)	1.07 (0.68, 1.71)		1.03 (0.58, 1.83)	
Shallow well water (less than 100ft)	0.66 (0.37, 1.18)		0.47 (0.21, 1.05)	
Other sources	1.17 (0.75, 1.83)		1.27 (0.76, 2.14)	
Fuel source – Natural gas		0.27		0.11
Yes	1.25 (0.84, 1.86)		1.49 (0.91, 2.41)	
No	1.00		1.00	
Household Propane Use		**0.18**		0.14
Yes	0.54 (0.22, 1.33)		0.42 (0.13, 1.33)	
No	1.00		1.00	
Mildew odor or musty smell in home		**0.14**		0.09
Yes	0.67 (0.40, 1.14)		0.54 (0.27, 1.10)	
No	1.00		1.00	
INDIVIDUAL FACTORS				
Smoking Status		0.35		0.90
Current Smoker	0.62 (0.32, 1.21)		0.88 (0.42, 1.88)	
Ex-smoker	0.88 (0.60, 1.30)		0.92 (0.58, 1.44)	
Never smoker	1.00		1.00	
Alcohol consumption		**0.10**		0.44

Never	1.00		1.00	
Less than once a month	0.93 (0.59, 1.47)		0.73 (0.41, 1.30)	
At most 2-3 times a month	0.53 (0.31, 0.89)		0.58 (0.32, 1.05)	
At most 2-3 times a week	0.62 (0.36, 1.09)		0.73 (0.39, 1.36)	
Everyday	0.79 (0.37, 1.66)		0.58 (0.25, 1.37)	
Physical activity		0.86		0.44
Yes	1.03 (0.71, 1.51)		0.85 (0.58, 1.29)	
No	1.00		1.00	
Early life-exposures:				
Ever lived on a farm		**0.13**		0.34
Yes	1.49 (0.89, 2.50)		1.38 (0.71, 2.69)	
No	1.00		1.00	
Lived on a farm in first year of life		**0.01**		0.20
Yes	1.74 (1.17, 2.59)		1.35 (0.85, 2.14)	
No	1.00		1.00	
Familial History of cancer:				
Father ever had cancer		**0.08**		0.43
Yes	1.40 (0.95, 2.02)		1.21 (0.76, 1.92)	
No	1.00		1.00	
Mother ever had cancer		**0.03**		0.58
Yes	1.52 (1.03, 2.23)		1.14 (0.71, 1.83)	
No	1.00		1.00	
Sibling(s) ever had cancer		**<0.001**		**<0.001**
Yes	3.28 (2.25, 4.76)		2.28 (1.44, 3.62)	
No	1.00		1.00	
Occupational Exposures:				
At work, ever exposed to:				
Asbestos dust	1.90 (0.87, 4.17)	**0.11**	0.52 (0.13, 2.15)	0.37
Diesel fumes	0.99 (0.69, 1.44)	0.97	0.94 (0.61, 1.44)	0.76
Fungicides (to treat grain)	1.28 (0.82, 2.00)	0.28	1.24 (0.76, 2.03)	0.40
Grain dust	1.25 (0.87, 1.79)	**0.23**	1.31 (0.84, 2.04)	0.23
Pesticides (to kill plants and insects)	1.52 (1.05, 2.19)	**0.03**	1.56 (1.03, 2.36)	0.04
Livestock	1.06 (0.74, 1.52)	0.76	1.04 (0.68, 1.58)	0.87
Mine dust	0.44 (0.06, 3.17)	0.41	0.48 (0.07, 3.55)	0.48
Molds	0.97 (0.65, 1.45)	0.88	1.04 (0.67, 1.64)	0.86
Oil/Gas well fumes	0.91 (0.53, 1.57)	0.72	0.99 (0.55, 1.78)	0.98

Radiation	4.04 (2.63, 6.21)	**<0.001**	3.94 (2.44, 6.36)	**<0.001**
Stubble smoke	1.14 (0.78, 1.67)	0.50	1.20 (0.78, 1.89)	0.41
Solvent fumes	0.91 (0.58, 1.43)	0.68	0.76 (0.44, 1.30)	0.32
Welding fumes	0.82 (0.47, 1.43)	0.48	0.91 (0.49, 1.68)	0.75
Wood dust	1.02 (0.67, 1.55)	0.91	1.06 (0.67, 1.68)	0.81
Covariates				
Age (yrs.)		**<0.001**		**<0.001**
18-45	1.00		1.00	
46-55	5.01 (1.45, 17.38)		2.06 (0.41, 10.24)	
56-65	13.27 (4.07, 43.22)		4.26 (0.97, 18.61)	
>=65	25.02 (7.87, 79.57)		14.78 (3.61, 60.61)	
BMI (kg/m^2)		0.44		0.66
Normal (0-<25)	1.00		1.00	
Overweight (25-30)	1.31 (0.86, 1.99)		1.17 (0.72, 1.89)	
Obese (>30)	1.15 (0.72, 1.82)		0.91 (0.53, 1.58)	
Education		0.05		0.18
≤ Grade 12	1.45 (1.01, 2.09)		1.33 (0.88, 2.01)	
> Grade 12	1.00		1.00	
Marital status		0.002		0.11
Married/common law/living together			0.68 (0.42, 1.10)	
Widowed/Divorced/separated/single/never married	1.00		1.00	

4.4.2.4 Multivariable analysis of Breast Cancer

Table 4.15-A presents the multivariable associations between breast cancer and predictors at both baseline and follow-up. At baseline, after a multivariable adjustment, statistically significant predictors of breast cancer included; sibling's positive history of other malignancies, work-related exposure to radiation, and age while that of follow-up included; work-related exposure to radiation, and age. These were associated with higher odds of developing breast cancer among persons with such characteristics.

In particular, a statistically significant dose-response relation was observed between the odds of developing breast cancer and age at baseline although the confidence intervals associated with the odds ratios appear to be fairly wide (Table 4.15-A).

TABLE 4.15-A: Multivariable Analysis of the Relationship between Breast cancers and Contextual Factors, Individual Factors, and Covariates by Odds Ratio (OR), 95% CI, and P-value by Time-point

Predictor	**Baseline (2010)**			**Follow-up (2014)**		
Description	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Unadjusted Odds Ratio (OR)* (95% CI)	P-value	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Unadjusted Odds Ratio (OR)* (95% CI)	P-value
CONTEXTUAL FACTORS						
Environmental:						
Water source						0.29
Bottled Water		1.00			1.00	
Deep well water				-1.39 (1.02)	0.25 (0.03, 1.85)	
Shallow well water				-0.86 (0.49)	0.42 (0.16, 1.11)	
Other sources				-0.11 (0.31)	0.89 (0.48, 1.66)	
Household Propane Use						0.17
Yes				-1.39 (1.02)	0.25 (0.03, 1.85)	
No		1.00			1.00	
Mildew odor or musty smell in home			0.20			
Yes	-0.39 (0.34)	0.68 (0.35, 1.33)		-0.63 (0.45)	0.54 (0.22, 1.29)	0.16
No		1.00			1.00	
Familial History of cancer:						
Father ever had cancer			0.20			

Yes	0.20 (0.22)	1.23 (0.80, 1.87)				
No		1.00			1.00	
Sibling(s) ever had cancer			0.001			
Yes	0.72 (0.22)	2.05 (1.34, 3.13)				
No		1.00			1.00	
Occupational Exposures:						
At work, ever exposed to:						
Grain dust				0.16 (0.28)	1.17 (0.66, 2.05)	0.58
Radiation	1.37 (0.26)	3.94 (2.41, 6.43)	<0.001	1.33 (0.30)	3.80 (2.11, 6.82)	<0.001
Covariates						
Age (yrs.)			<0.001			<0.001
18-45		1.00			1.00	
46-55	1.37 (0.65)	3.95 (1.11, 14.01)		1.14 (1.12)	3.13 (0.35, 28.10)	
56-65	2.30 (0.61)	9.97 (3.04, 32.74)		1.99 (1.05)	6.99 (0.89, 24.55)	
>=65	2.72 (0.61)	15.18 (4.62, 49.83)		3.2 (1.01)	24.64 (3.38,79.77)	
BMI (kg/m^2)			0.86			0.83
Normal (0-<25)		1.00			1.00	
Overweight (25-30)	0.21 (0.25)	1.24 (0.76, 2.01)		-0.01 (0.30)	0.99 (0.55, 1.78)	
Obese (>30)	0.14 (0.27)	1.15 (0.68, 1.94)		-0.19 (0.34)	0.83 (0.42, 1.61)	
Education						0.60
≤ Grade 12				0.15 (0.27)	1.16 (0.68, 1.97)	
> Grade 12		1.00			1.00	
Marital status						0.69
Married/common law/living together				0.11 (0.28)	1.12 (0.64, 1.94)	

Widowed/Divorced/separated/single/never married	1.00	1.00

As mentioned earlier in this thesis that, for the purposes of prediction probabilities for breast cancer at the baseline and follow up, we repeated the same steps of model building as carried out for HRCs multivariate analysis. On obtaining the final model, it was found that the following variables (i.e. history of sibling cancer, exposure to radiations, and age) at the baseline independently contributed to increased risk of breast cancer (Table 4.15-B). However, at the follow up our final model revealed that the significant predictor of breast cancer included; quadrant, exposure to radiation, and age. (Table 4.15- B)

Table 4.15-B: Predictor probabilities; Multivariable Analysis of the Relationship between Breast cancers and Contextual Factors, Individual Factors, and Covariates by Odds Ratio (OR), 95% CI, and P-value by Time-point

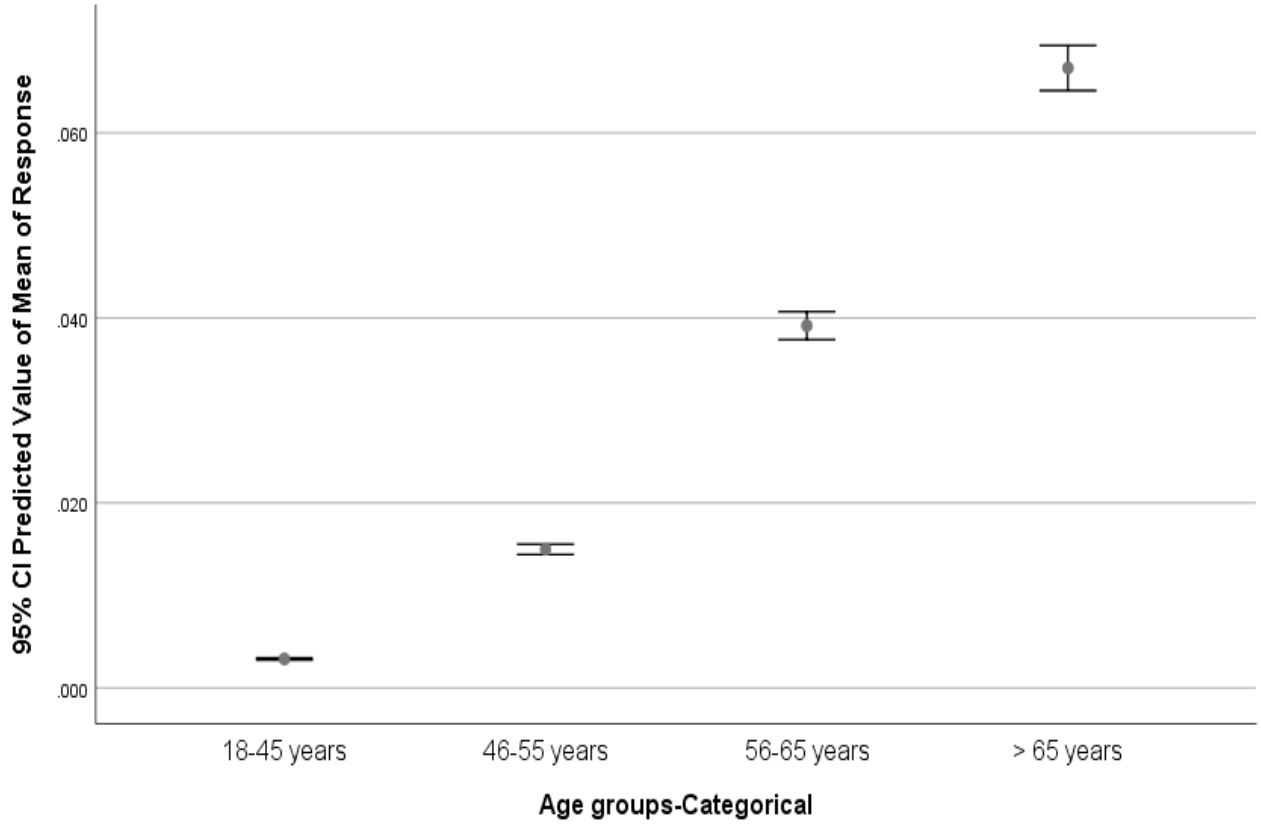
Predictor		**Baseline (2010)**			**Follow-up (2014)**	
Description	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Adjusted Odds Ratio (OR)* (95% CI)	P-value	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Adjusted Odds Ratio (OR)* (95% CI)	P-value
CONTEXTUAL FACTORS						
Location of home						
Farm	-0.07 (0.21)	0.93 (0.61,1.40)	0.714	-0.21 (0.24)	0.80 (0.51, 1.29)	0.372
Sibling(s) ever had cancer						
Yes	0.65 (0.20)	1.92 (1.29, 2.87)	**<0.001**		----------	------
Occupational Exposures:						
At work, ever exposed to:						
Radiation	1.26 (0.24)	3.52 (2.19, 5.66)	**<0.001**	1.33 (0.25)	3.77 (2.31, 6.14)	**<0.001**
Covariates						
Age (yrs.)						

18-45	Ref.	1.00		Ref.	1.00	
46-55	1.51 (0.64)	4.55 (1.29, 15.9)	**0.018**	0.77 (0.83)	2.15 (0.43, 10.9)	0.353
56-65	2.42 (0.60)	11.29 (3.5, 36.6)	**<0.001**	1.34 (0.76)	3.8 (0.86, 16.9)	0.079
>=65	2.85 (0.60)	17.35 (5.3, 56.6)	**<0.001**	2.60 (0.73)	1.35 (3.2, 56.7)	**<0.001**
Quadrant						
South West	0	1		0	1	
South East	0.04 (0.31)	1.05 (0.57, 1.93)	0.883	0.87 (0.39)	2.39 (1.10, 5.19)	**0.028**
North East	0.06 (0.28)	0.84 (0.47, 1.51)	0.836	0.59 (0.39)	1.82 (0.84, 3.94)	0.154
North West	-0.17 (0.29)	1.06 (0.60, 1.87)	0.570	0.55 (0.38)	1.74 (0.81, 3.73)	0.130

***-----** indicates that the corresponding variables were not found significant at the final model although we retained the purposeful or otherwise biologically relevant variables in the model.

The probability of getting breast cancer with age- and work-related exposure to radiations at the baseline are shown below in the figure 4.3a, and graph 4.3b, respectively.

Figure 4.3a Breast cancer and Age (Baseline)



The figure 4.3a above shows the probability of getting breast cancer with age at the baseline.

Overall a clear dose-response probability of having breast cancer with increase in the age was observed (i.e. 0.31% to 6.70%).

In terms of discrete age groups, the probability of having breast cancer was 0.31% in the age group of 18-45 years, however for age group 46-55 years the probability constituted 1.50%, whereas for age group 56-65 years, the probability was 3.91%, and finally for age group >65 years this reached to 6.70 % (Figure 4.3a)

Figure 4.3b: Breast cancer and Exposure to Radiations (Baseline)

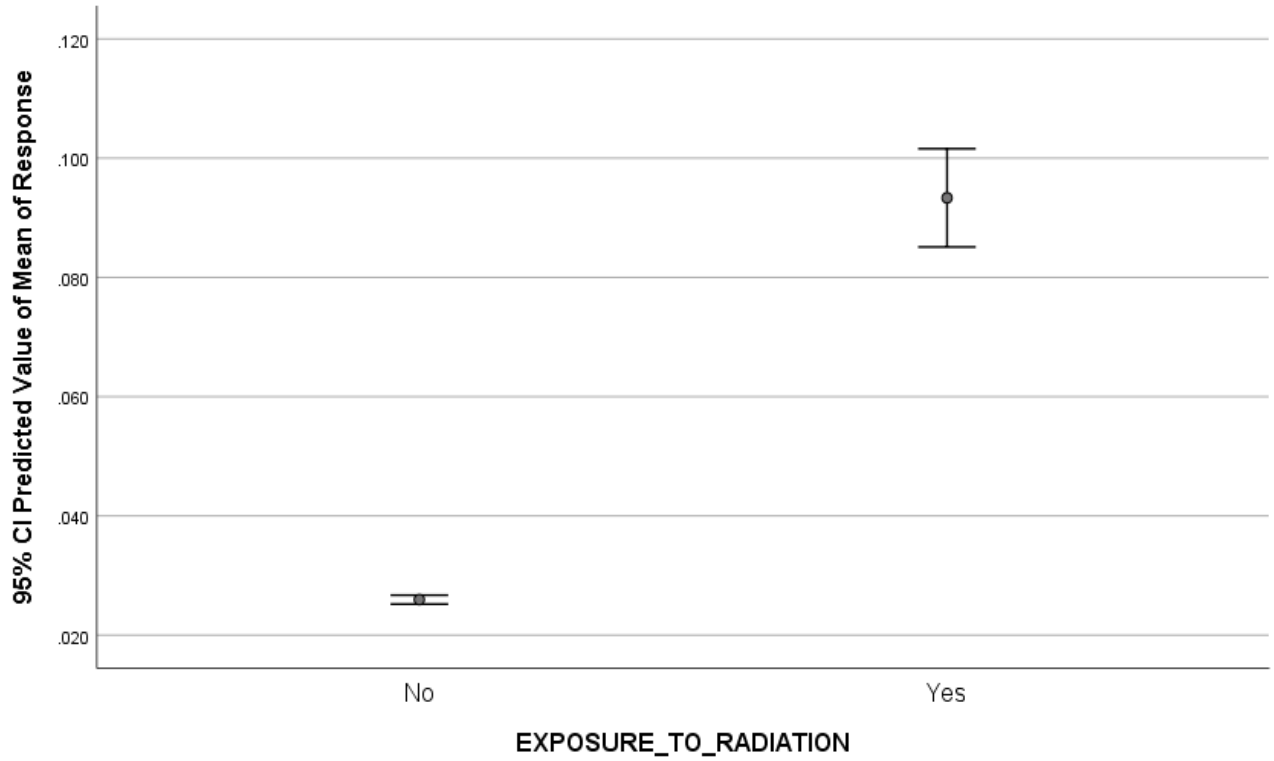
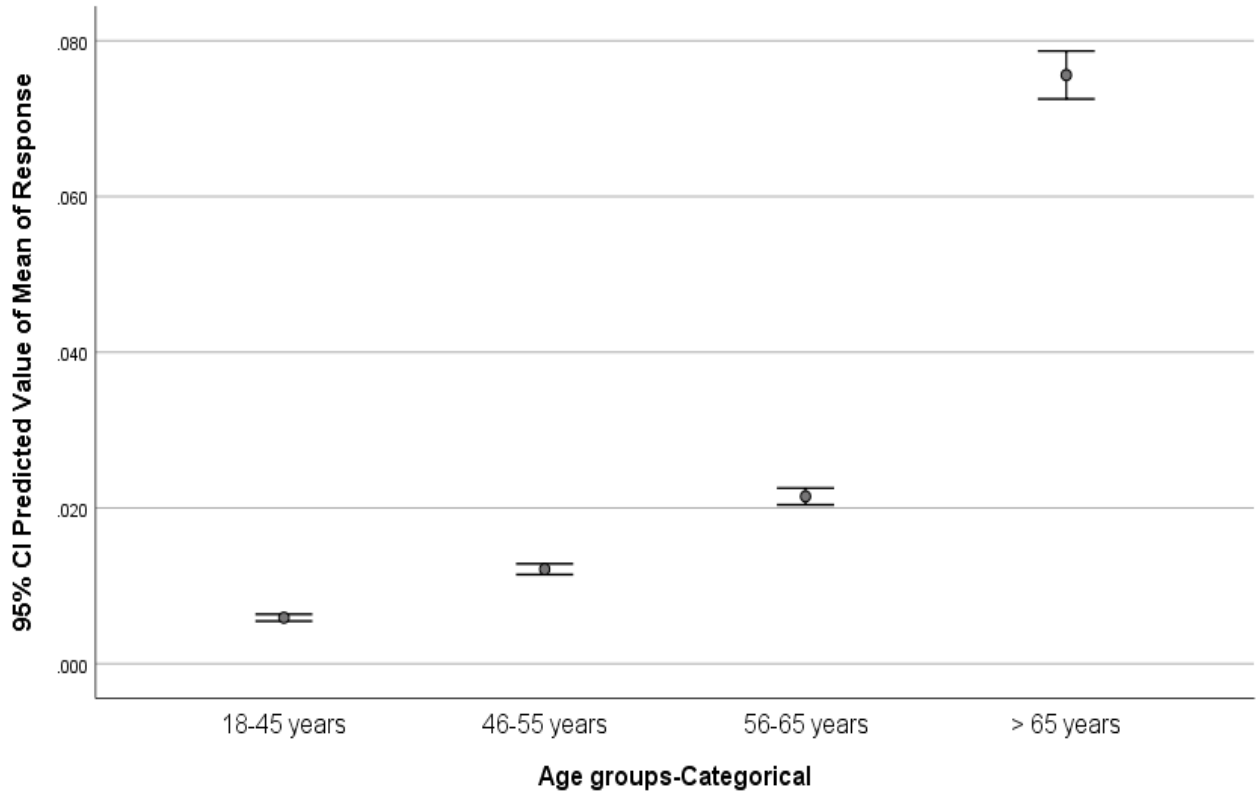


Figure 4.3b above shows the probability of getting breast cancer with exposure to radiation at the baseline. It is evident from the figure 4.3b, that a probability of getting breast cancer among those exposed to radiations was greater (9.33%) than those who were not exposed (2.60%).

Now, the probability of getting breast cancer with age, and work-related exposure to radiations at the follow up are shown below in the figure 4.4a, and figure 4.4b, respectively.

Figure 4.4a Breast cancer and Age (Follow-up)



The figure 4.4a shows the probability of getting breast cancer with age at the follow-up. Overall a clear dose-response probability of having breast cancer with increase in the age was observed (i.e. 0.59% to 7.56%).

Considering various age groups, the probability of having breast cancer was 0.59% in the age group of 18-45 years, however for the age group 46-55 years; the probability comprised 1.21%, whereas for the age group 56-65 years; the probability of 2.15%, was observed. Finally, this attained 7.56 % for age group >65 years (Figure 4.4a)

Figure 4.4b Breast cancer and Exposure to Radiations (Follow-up)

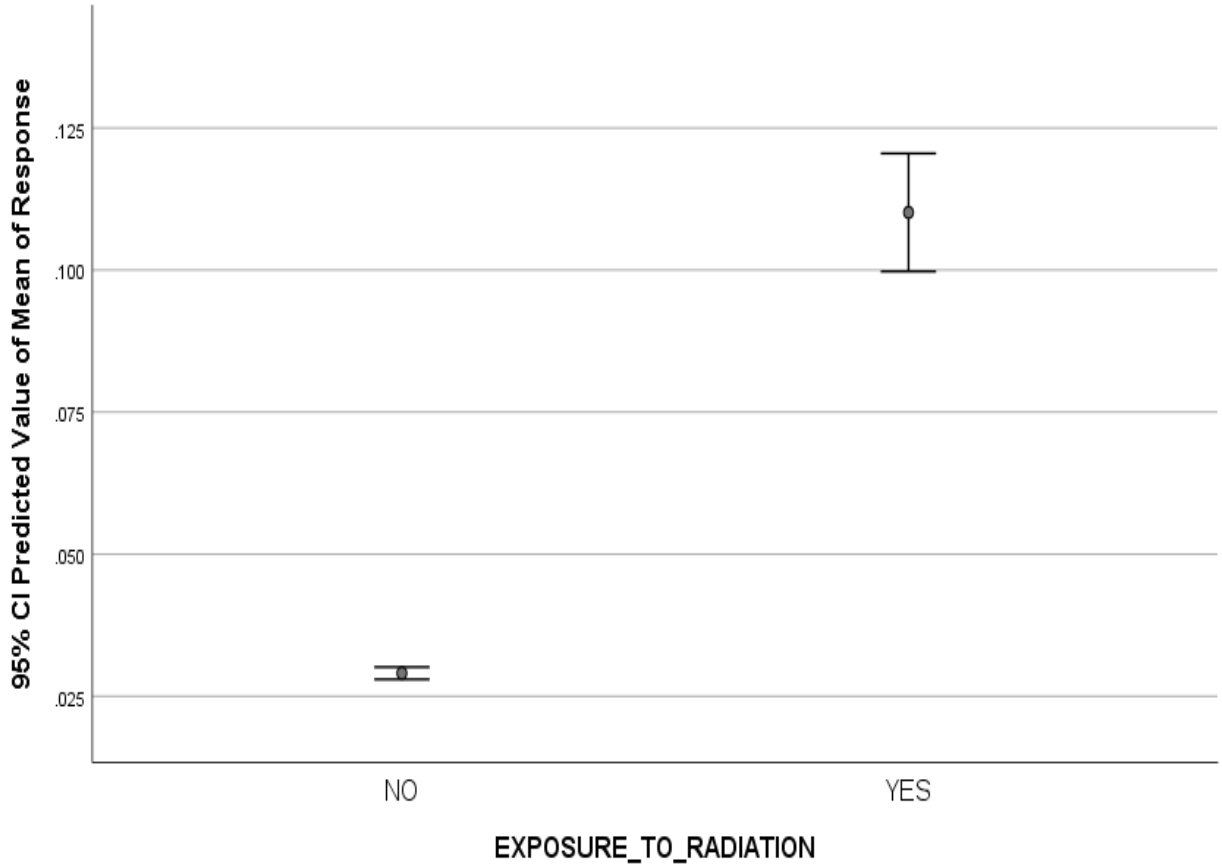


Figure 4.4b above shows the probability of getting breast cancer with work related exposure to radiation at the follow up.

It is evident from the figure above, that a probability of having breast cancer among those exposed to radiations was greater (11.01%) than those who were not exposed (2.90%).

4.4.2.5 Univariable analysis of Prostate Cancer:

This section presents the univariable associations between prostate cancer and risk factors both at baseline and follow-up. With a p-value<0.25, significant predictors of prostate cancer in the univariable analysis at baseline were; quadrant, household smoking, household water source, mildew or musty smell in home, smoking status, alcohol consumption, ever living on farm, living on farm in first year of life, family history of cancer (i.e. father, mother, and sibling's positive history of cancer), occupational exposure to asbestos dust, fungicides, grain dust, pesticides,

livestock, mine dust, radiation, smokes stubble, welding fumes, BMI, and educational status (Table 4.16).

On the other, hand, significant predictors of prostate cancers at follow-up in the univariable analysis included; quadrant, household use of natural gas, smoking status, alcohol consumption, ever living on farm, sibling's positive history of cancer, diesel fumes, grain dust, pesticides, mine dust, oil and gas well fumes, radiation, welding fumes, and educational status (Table 4.16)

TABLE 4.16 Univariable Analysis of the Relationship between Prostate cancer and Contextual Factors, Individual Factors, and Covariates by Odds Ratio (OR), 95% CI, and P-value by Time-point

Predictor	**Baseline (2010)**		**Follow-up (2014)**	
Description	Unadjusted Odds Ratio (OR)* (95% CI)	P-value[#]	Unadjusted Odds Ratio (OR)* (95% CI)	P-value[#]
CONTEXTUAL FACTORS				
Socioeconomic		0.59		0.39
Some money	1.00		1.00	
Just enough money	0.76 (0.44, 1.31)		1.23 (0.69, 2.20)	
Not enough money	0.87 (0.51, 1.48)		0.61 (0.24, 1.55)	
Quadrant		**0.13**		**0.18**
South West	1.00		1.00	
South East	0.95 (0.55, 1.67)		0.75 (0.39, 1.47)	
North East	1.10 (0.66, 1.83)		0.66 (0.35, 1.25)	
North West	0.60 (0.34, 1.07)		0.47 (2.4, 0.93)	
Environmental:				
Household smoking		**0.02**		0.53
Yes	0.44 (0.22, 0.88)		0.76 (0.33, 1.77)	
No	1.00		1.00	
Location of home		0.93		0.73
Farm	1.02 (0.70, 1.48)		0.92 (0.56, 1.47)	
Non-farm	1.00		1.00	
Water source		**0.03**		0.94
Bottled Water	1.00		1.00	

Deep well water (more than 100ft)	1.19 (0.71, 1.99)		1.12 (0.59, 2.14)	
Shallow well water (less than 100ft)	0.64 (0.34, 1.22)		1.04 (0.53, 2.06)	
Other sources	1.56 (0.97, 2.53)		0.91 (0.48, 1.73)	
Fuel source – Natural gas		0.81		**0.25**
Yes	0.95 (0.65, 1.41)		1.36 (0.81, 2.30)	
No	1.00		1.00	
Household Propane Use		0.98		0.64
Yes	0.99 (0.50, 1.98)		0.80 (0.32, 2.02)	
No	1.00		1.00	
Mildew odor or musty smell in home		**0.22**		0.27
Yes	1.33 (0.84, 2.11)		0.66 (0.31, 1.38)	
No	1.00		1.00	
INDIVIDUAL FACTORS				
Smoking Status		**0.001**		**0.03**
Current Smoker	0.89 (0.45, 1.78)		0.80 (0.28, 2.33)	
Ex-smoker	1.75 (1.19, 2.59)		1.82 (1.11, 2.98)	
Never smoker	1.00		1.00	
Alcohol consumption		**<0.001**		**0.04**
Never	1.00		1.00	
Less than once a month	0.53 (0.30, 0.93)		0.40 (0.17, 0.92)	
At most 2-3 times a month	0.51 (0.31, 0.83)		0.58 (0.30, 1.11)	
At most 2-3 times a week	0.29 (0.17, 0.51)		0.34 (0.16, 0.69)	
Everyday	0.35 (0.18, 0.68)		0.64 (0.32, 1.30)	
Physical activity		0.68		0.98
Yes	1.08 (0.74, 1.58)		1.01 (0.63, 1.61)	
No	1.00		1.00	
Early life-exposures:				
Ever lived on a farm		**0.002**		0.64
Yes	3.61 (1.58, 8.25)		1.19 (0.57, 2.51)	
No	1.00		1.00	
Lived on a farm in first year of life		**0.008**		0.59
Yes	2.02 (1.20, 3.40)		1.18 (0.64, 2.18)	
No	1.00		1.00	
Familial History of cancer:				
Father ever had cancer		**0.003**		**0.002**
Yes	1.83 (1.23, 2.72)		2.24 (1.33, 3.77)	
No	1.00		1.00	

Mother ever had cancer		**0.08**		0.26
Yes	1.47 (0.96, 2.26)		1.37 (0.79, 2.38)	
No	1.00		1.00	
Sibling(s) ever had cancer		**<0.001**		**<0.001**
Yes	2.52 (1.67, 3.79)		2.91 (1.64, 5.15)	
No	1.00		1.00	
Occupational Exposures:				
At work, ever exposed to:				
Asbestos dust	1.00 (0.55, 1.84)	0.99	1.39 (0.72, 2.68)	0.32
Diesel fumes	1.18 (0.73, 1.90)	0.51	1.99 (0.86, 4.61)	**0.11**
Fungicides (to treat grain)	1.86 (1.27, 2.73)	**0.002**	1.31 (0.80, 2.14)	0.29
Grain dust	1.71 (0.93, 3.12)	**0.08**	2.33 (0.84, 6.43)	**0.10**
Pesticides (to kill plants and insects)	2.13 (1.42, 3.17)	**<0.001**	1.50 (0.80, 2.81)	**0.20**
Livestock	1.70 (1.12, 2.57)	**0.01**	1.34 (0.77, 2.32)	0.30
Mine dust	0.53 (0.23, 1.21)	**0.13**	0.32 (0.10, 1.03)	**0.06**
Molds	1.01 (0.69, 1.46)	0.98	0.91 (0.57, 1.46)	0.69
Oil/Gas well fumes	1.14 (0.78, 1.67)	0.50	1.48 (0.93, 2.37)	**0.10**
Radiation	2.40 (1.49, 3.87)	**<0.001**	2.28 (1.27, 4.09)	**0.01**
Stubble smoke	1.42 (0.98, 2.08)	**0.06**	1.66 (1.00, 2.76)	**0.05**
Solvent fumes	0.84 (0.58, 1.21)	0.34	0.96 (0.60, 1.54)	0.87
Welding fumes	1.58 (1.01, 2.45)	**0.04**	1.41 (0.78, 2.55)	**0.25**
Wood dust	0.94 (0.65, 1.36)	0.75	1.22 (0.75, 1.99)	0.43
Covariates				
BMI (kg/m^2)		**0.03**		0.66
Normal (0-<25)	1.00		1.00	
Overweight (25-30)	0.64 (0.41, 1.00)		0.76 (0.40, 1.42)	
Obese (>30)	0.52 (0.31, 0.86)		0.89 (0.46, 1.73)	
Education		**0.001**		**0.05**
≤ Grade 12	2.33 (1.42, 3.83)		1.75 (0.99, 3.07)	
> Grade 12	1.00		1.00	
Marital status		0.26		0.83
Married/common law/living together	0.76 (0.48, 1.21)		1.09 (0.52, 2.28)	
Widowed/Divorced/separated/single/never married	1.00		1.00	

4.4.2.6 Multivariable analysis of Prostate Cancer

Variables selected in section 4.4.2.5 as significant were candidates for the multivariable modelling. Table 4.17-A presents the results of the multivariable associations between prostate cancer and predictors at both baseline and follow-up.

At baseline, statistically significant predictors of prostate cancer included; sibling's positive history of other malignancies, work-related exposure to radiation or pesticides, and age.

On the other hand, significant predictors of prostate cancer at follow-up included household use of natural gas and sibling's positive history of cancer (Table 4.17-A).

TABLE 4.17-A: Multivariable Analysis of the Relationship between Prostate cancers and Contextual Factors, Individual Factors, and Covariates by Odds Ratio (OR), 95% CI, and P-value by Time-point

Predictor	**Baseline (2010)**			**Follow-up (2014)**		
Description	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Unadjusted Odds Ratio (OR)* (95% CI)	P-value	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Unadjusted Odds Ratio (OR)* (95% CI)	P-value
CONTEXTUAL FACTORS						
Quadrant			0.11			0.06
South West		1.00			1.00	
South East	-0.46 (0.35)	0.63 (0.32, 1.27)		-1.30 (0.56)	0.27 (0.09, 0.81)	
North East	-0.37 (0.31)	0.69 (0.38, 1.28)		-0.69 (0.42)	0.50 (0.22, 1.15	
North West	-0.88 (0.36)	0.42 (0.21, 0.84)		-1.05 (0.48)	0.35 (0.13, 0.91)	
Environmental:						
Household smoking			0.06			
Yes	-0.90 (0.48)	0.41 (0.16, 1.04)				
No		1.00			1.00	
Water source			0.06			
Bottled Water		1.00				

Deep well water	-0.08 (0.32)	0.92 (0.49, 1.73)			
Shallow well water	-0.66 (0.39)	0.52 (0.24, 1.11)			
Other sources	0.36 (0.30)	1.44 (0.80, 2.58)			
Fuel source – Natural gas					0.04
Yes			0.85 (0.41)	2.33 (1.05, 5.18)	
No				1.00	
Mildew odor or musty smell in home			0.11		
Yes	0.42 (0.26)	1.52 (0.91, 2.56)			
No		1.00		1.00	
INDIVIDUAL FACTORS					
Smoking Status			0.14		
Current Smoker	0.46 (0.48)	1.58 (0.62, 4.03)			
Ex-smoker	0.50 (0.26)	1.65 (1.00, 2.73)			
Never smoker		1.00		1.00	
Alcohol consumption			0.05		0.15
Never		1.00		1.00	
Less than once a month	-0.31 (0.35)	0.73 (0.37, 1.45)	-1.13 (0.60)	0.32 (0.10, 1.04)	
At most 2-3 times a month	-0.53 (0.33)	0.59 (0.31, 1.12)	-0.82 (0.52)	0.44 (0.16, 1.22)	
At most 2-3 times a week	-0.99 (0.36)	0.37 (0.18, 0.76)	-1.26 (0.54)	0.28 (0.10, 0.81)	
Everyday	-0.99 (0.44)	0.37 (0.16, 0.88)	-0.56 (0.50)	0.57 (0.21, 1.53)	
Early life-exposures:					
Ever lived on a farm			0.15		
Yes	0.72 (0.49)	2.05 (0.78, 5.35)			
No		1.00		1.00	

Familial History of cancer:						
Father ever had cancer			0.07			0.40
Yes	0.44 (0.25)	1.56 (0.96, 2.53)		0.28 (0.33)	1.33 (0.69)	
No		1.00			1.00	
Sibling(s) ever had cancer			0.02			0.01
Yes	0.58 (0.25)	1.78 (1.08, 2.93)		0.92 (0.33)	2.52 (1.31, 4.84)	
No		1.00			1.00	
Occupational Exposures:						
At work, ever exposed to:						
Grain dust				0.76 (0.89)	2.13 (0.37, 12.12)	0.39
Pesticides (to kill plants and insects)	0.58 (0.28)	1.80 (1.04, 3.11)	0.04			
Livestock	0.22 (0.30)	1.24 (0.68, 2.26)	0.48			
Mine dust	-0.84 (0.55)	0.43 (0.15, 1.26)	0.12	-0.53 (0.66)	0.59 (0.16, 2.16)	0.43
Molds						
Oil/Gas well fumes						
Radiation	0.73 (0.30)	2.08 (1.15, 3.77)	0.02	0.71 (0.46)	2.03 (0.83, 4.99)	0.12
Stubble smoke				0.41 (0.46)	1.51 (0.61, 3.70)	0.37
Solvent fumes						
Welding fumes				-0.31 (0.46)	0.74 (0.30, 1.80)	0.50
Covariates						
BMI (kg/m^2)			0.15			0.51
Normal (0-<25)		1.00			1.00	
Overweight (25-30)	-0.42 (0.28)	0.66 (0.38, 1.13)		0.53 (0.47)	1.70 (0.68, 4.29)	
Obese (>30)	-0.56 (0.32)	0.56 (0.30, 1.05)		0.53 (0.47)	1.67 (0.61, 4.57)	
Education			0.06			0.18

≤ Grade 12	0.61 (0.33)	1.84 (0.97, 3.51)		0.51 (0.38)	1.67 (0.79, 3.51)
> Grade 12		1.00			1.00
Marital status					
Married/common law/living together					
Widowed/Divorced/separated/single /never married		1.00			1.00

For the purposes of prediction probabilities for prostate cancer at the baseline and follow up, we adopted the standard model building approach as it was employed earlier to carry out univariable and multivariable analysis for HRCs and breast cancer.

On obtaining the final model, it was found that the following variables were significant in the multivariate analysis at the baseline; age, work related exposure to radiations and pesticides, alcohol consumption, household water source, and musty smell (Table 4.17-B).

However, at the follow up, the significant predictor of prostate cancer in the multivariate analysis included age only. (Table 4.17-B)

It is imperative to note that the age categories used previously in model building for HRCs and breast cancer were not employed in model building for assessing univariate and multivariate association between prostate cancer and risk factors, both at the baseline and follow up.

Instead we collapsed the age category of (18 years – 45 years) and (46 years to 55 years). This was because the number of cases in either category (i.e. 18 – 45 years) and (35- 55 years) were low and thus was producing unstable results (i.e. wide CIs).

TABLE 4.17-B: Predictor Probabilities; Multivariable Analysis of the Relationship between Prostate cancers and Contextual Factors, Individual Factors, and Covariates by Odds Ratio (OR), 95% CI, and P-value by Time-point

Predictor	**Baseline (2010)**			**Follow-up (2014)**		
Description	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Unadjusted Odds Ratio (OR)* (95% CI)	P-value	$\widehat{\boldsymbol{\beta}}[\boldsymbol{SE}(\widehat{\boldsymbol{\beta}})]$	Unadjusted Odds Ratio (OR)* (95% CI)	P-value

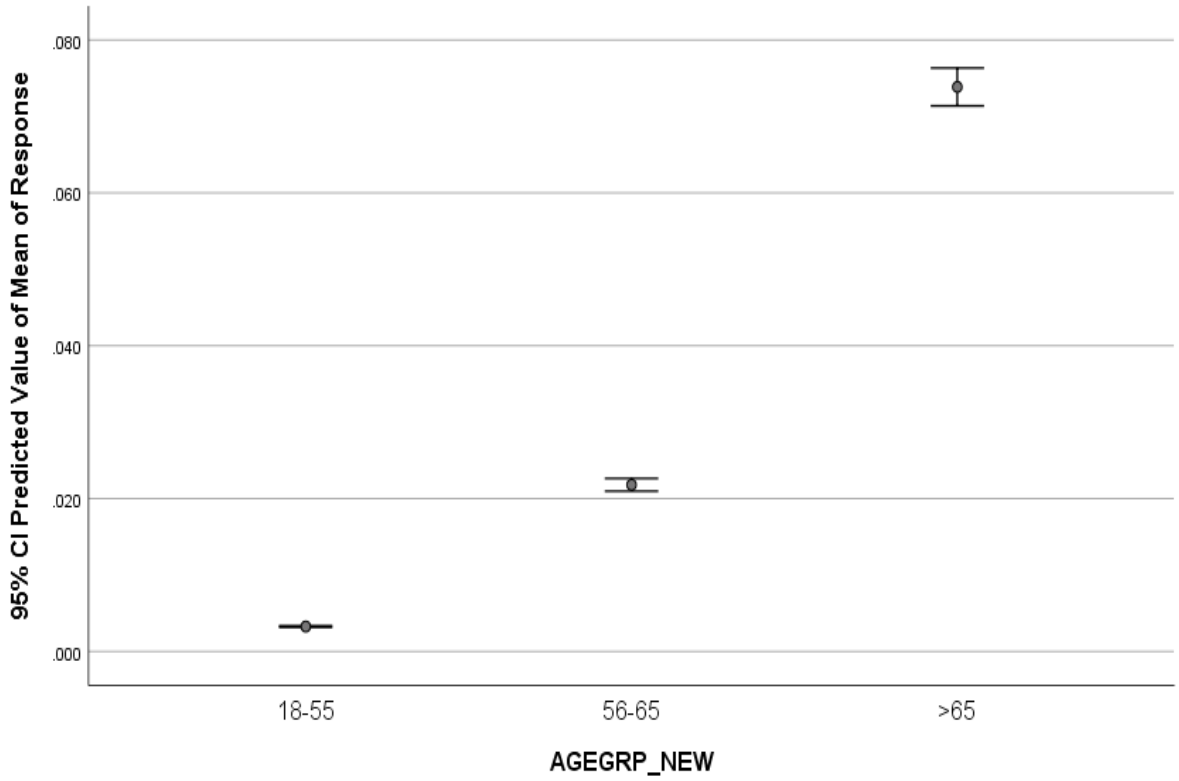
CONTEXTUAL FACTORS						
Quadrant						
South West		1.00				
South East	-0.19 (0.32)	0.83 (0.44, 1.54)	0.55	-0.70 (0.53)	0.49 (0.17, 1.39)	0.18
North East	-1.70 (0.27)	0.84 (0.49, 1.45)	0.54	-0.37 (0.42)	0.68 (0.30, 1.56)	0.37
North West	-0.43 (0.31)	0.65 (0.35, 1.18)	0.16	-0.71 (0.47)	0.49 (0.19, 1.24)	0.13
Environmental:						
Location of home						
Farm	0.17 (0.22)	1.19 (0.77, 1.85)	0.43	-0.34 (0.34)	0.71 (0.37, 1.38)	0.32
Water source						
Bottled water		1.00				
Deep well water (more than 100 ft)	-0.12 (0.28)	0.98 (0.57, 1.71)	0.95 **0.02**	-------- --------	-------- --------	------ --
Shallow well water (less than 100 ft)	-0.78 (0.35)	0.45 (0.23, 0.90)	0.39	-------- --------	-------- --------	------ --
Other sources	0.22 (0.26)	1.25 (0.001, 0.01)				------ --
Mildew odour or musty smell in home						
Yes	0.54 (0.25)	1.72 (1.06, 2.80)	**0.03**	--------	--------	------ --
INDIVIDUAL FACTORS						
Alcohol consumption						
Never	-0.29 (0.31)	0.75 (0.41, 1.37)	0.35	------	----------	------ -
Less than once a month	-0.13 (0.27)	0.88 (0.52, 1.51)	0.65	------	----------	------ -
At most 2-3 times a month	- 0.48(0.31)	0.61 (0.33, 1.13)	0.12 **0.05**	------ ------	---------- ----------	------ -
Everyday	-0.71(0.37)					

		0.50 (0.24, 1.02)				------
Occupational Exposures:						
At work, ever exposed to:						
Pesticides (to kill plants and insects)	0.48 (0.22)	1.61 (1.04, 2.50)	**0.03**	-------	----------	------
Radiation	0.67 (0.26)	1.96 (1.17, 3.29)	**0.01**	-------	----------	------
Covariates						
Age (yrs)						
18 – 55		1.00			1.00	
56- 65	1.94 (0.46)	6.94 (2.8, 17.2)	**<0.00**	2.13	8.45 (1.05,	**0.04**
>65	3.2 (0.42)	23.4 (10.3,	**1**	(1.06)	67.8)	**<0.0**
		53.4)	**<0.00**	3.43	30.8 (4.2,	**01**
			1	(1.02)	227.4)	

***----** indicates that the corresponding variables were not found significant at the final model although we retained the purposeful or otherwise biologically relevant variables in the model.

Next, the probability of getting prostate cancer with age, and work-related exposure to radiations at the baseline are shown below in the figure 4.5a, and 4.5b, respectively.

Figure 4.5a: Prostate Cancer and Age (Baseline)



The figure 4.5a shows the probability of having prostate cancer with age at the baseline. Overall a clear dose-response probability of getting prostate cancer with increase in the age was observed (i.e. 0.32% to 7.38%).

In age groups like 18 -55 years, the probability of getting prostate cancer was 0.32%, however for the age group 56-65 years; the probability constituted 2.18%. Finally, this reached 7.38 % for age group >65 years (Figure 4.5a)

Figure 4.5b: Prostate Cancer and Exposure to Radiations (Baseline)

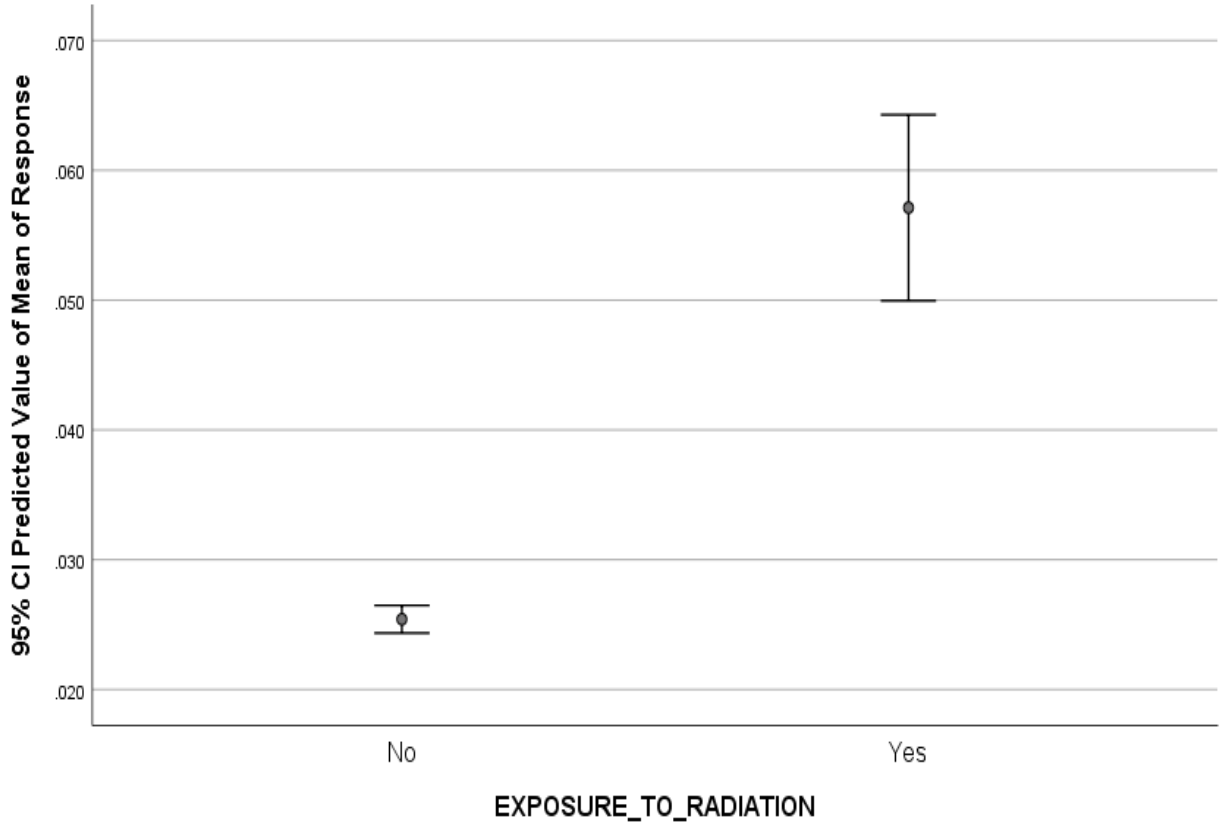


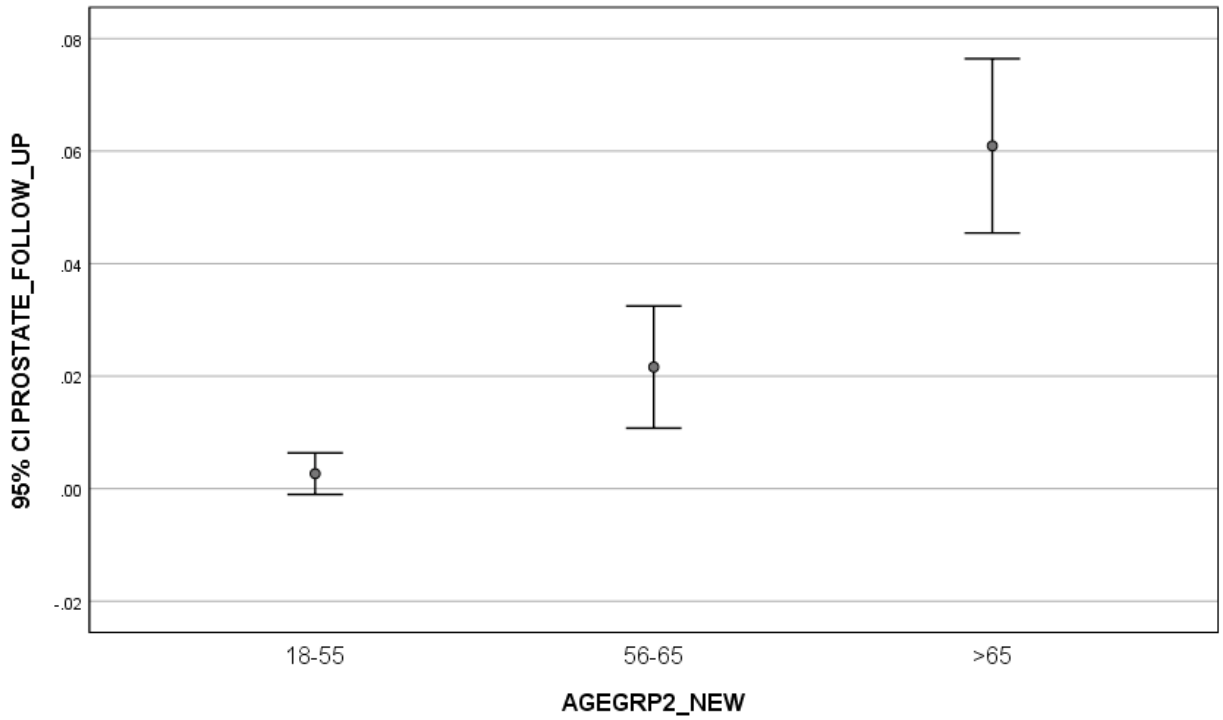
Figure 4.5b above shows the probability of getting prostate cancer with work-related exposure to radiation at the baseline.

It is evident from the figure above, that a probability of having prostate cancer among those exposed to radiations was greater (5.71%) than those who were not exposed (2.54%).

Here, a graphical presentation of age with prostate cancer at the follow-up is shown in figure 4.6a.

At this point, it is refreshing to note that the age groups used for prostate cancer were unlike those used for HRCs and breast cancer separately, i.e. for prostate cancer only three categories of age groups were employed: 18-55 years, 56-66 years, and >65 years. The reason was small numbers in the first two categories; therefore, a decision was made to combine the first two categories (i.e. 18-45, and 46-55 years).

Figure 4.6a Prostate Cancer and Age (Follow-up)



The figure 4.6a shows the probability of having prostate cancer with age at the follow up. Overall a clear dose-response probability of getting prostate cancer with increase in the age was observed (i.e. 0.22% to 6.55%).

In terms of discrete age groups like 18 -55 years, and 56-65 years, the probability of getting prostate cancer was 0.22%, and 1.81%, respectively. Although, the age group >65 years showed the highest probability (6.55%) (Figure 4.6a)

4.5 Incidence Analysis

4.5.1 Cumulative Incidence Calculation

I used the formula: Cumulative Incidence = (Total number of new cases in a specified period / Total number of people at risk in the specified period) * 100, to calculate the cumulative incidence for breast and prostate cancers combined and separately. Therefore, the incidence analysis contains only new cancer cases over the four years period (i.e. 2010 to 2014).

The table below (4.18) shows the calculation of the cumulative incidence for these malignancies.

Table 4.18 Calculation of New cases and Cumulative Incidence

Cancer type	Prevalence		Incidence [Method adopted from Ref (120)]	
	Baseline	Follow-up	Percentage of Valid Sample that must be New Cases (Cumulative Incidence)	Actual Number of New Cases
Breast cancer	$\frac{130}{4195} * 100 = 3.09\%$	$\frac{94}{2502} * 100 = 3.75\%$	$\frac{36}{4195} * 100 = 0.86\%$	36
Prostate	$\frac{117}{4065} * 100 = 2.87\%$	$\frac{73}{2364} * 100 = 3.08\%$	$\frac{44}{4065} * 100 = 1.08\%$	44
HRC	$\frac{247}{8260} * 100 = 2.99\%$	$\frac{167}{4866} * 100 = 3.43\%$	$\frac{80}{8260} * 100 = 0.97\%$	80

*From the 4741 individuals that took part in survey at two time-points, one male respondent who self-reported to be a case of breast cancer was excluded. This gave the total number of individuals that took part in both surveys as 4740.

From results contained in Table 4.18, there were 36 and 44 incident cases of breast and prostate cancers respectively after the four years of follow-up. In total, 80 incident cases of HRCs were observed. As a result, the cumulative incidence (incidence proportion) for breast and prostate cancers were 0.86% and 1.08%, respectively. When combined (i.e. breast and prostate cancer together) the cumulative incidence of HRCs was 0.97%. The number of disease-free individuals at the baseline for breast cancer were 4064 and at the follow-up were 2408. However, the number of disease-free individuals at the baseline for prostate cancer were 3947, and at follow-up were 2291.

Chapter 5: DISCUSSION & CONCLUSION

5.1 Introduction

In this chapter, the results of this study are discussed. Section 5.2 and its subsection 5.2.1 respectively, provides a discussion of the results for the prevalence (crude) and risk factors for the adjusted prevalence of HRCs. Incidence (crude) of HRCs are given in sections 5.3. Section 5.4 entails discussion around study strengths as well as the limitation of the current study under sub-sections 5.4.1 and 5.4.2, respectively. Finally, this study makes some recommendations in section 5.5 and stated the conclusion under section 5.6.

5.2 Prevalence

At the baseline, a total of 247 out of 8261 participants self-reported as doctor diagnosed HRC cases (i.e. 117 prostate and 130 breast cancer cases), however, at the follow-up, HRC cases were 167 out of 4867 (i.e. 94 breast and 73 prostate cancers cases). As mentioned earlier in this thesis that, the HRCs cases of 167 at the follow-up were inclusive of HRC cases at the baseline. In this study, the crude prevalence of HRC cases at the baseline and follow up were 3.0%, and 3.4%, respectively.

In particular, the unadjusted prevalence of prostate cancer was 3.0% and 3.3%, for farm and non-farm male residents respectively, at the follow-up. This finding of our study was comparable with a study conducted earlier by Sharma and colleagues on parts of rural Saskatchewan (115), where the age-standardized prevalence of prostate cancer was expressed as 3.32% (115).

A slight difference in the prevalence rates could be attributed to the selection of a different age group by Sharma and colleagues (i.e. 40 years and older) (115), however, the age group chosen for current project was 18 years and older.

For breast cancer, our study reported the crude prevalence of 3.0% and 4.2% among farm and non-farm female residents respectively, at the follow-up. To my knowledge, I did not find any study investigating the prevalence of breast cancer among female residents of rural Saskatchewan. Therefore, the unadjusted prevalence of breast cancer obtained in this analysis could not be compared for the province of Saskatchewan.

Generally, in Canada, the age-adjusted prevalence of cancers has increased slightly during the most recent 10 years (from 5.9% to 6.8%) (109). In particular, Ontario cancer statistics 2016 report identified prostate cancer as the leading 10-year prevalent cancer cases among males (1147 per

10^5), and for females, the leading prevalent cases were of carcinoma breast (988.2 per 10^5), during the years 2002-2012 (110).

While reviewing the studies that examined the relationship between farm residence and HRCs prevalence; Sharma and colleagues reported that farm residence was a significant risk factor that was associated with the prevalence of prostate cancer (115). Similarly, for breast cancers, an epidemiological study highlighted an association between breast cancer prevalence and living near agricultural areas (119).

Our study finding, however, contradicted this result as HRCs were slightly more prevalent amongst non-farm residents (3.2%) when compared with farm residents (2.8%).

5.2.1 Risk factors for the Adjusted Prevalence of HRCs

As described earlier in Chapter 4 that, after a multivariable adjustment, the significant predictors for the adjusted prevalence of HRCs at baseline were: the household source of water supply, father and sibling positive history of cancer, exposure to mine dust and radiation, and age.

For the household source of water supply, there is a growing body of epidemiological evidence expressing the association of household sources of water supply with one or more cancers, and this evidence has been consistent (101).

A study conducted in Iowa investigating the association between prostate cancer and level of arsenic in drinking water from public and private wells reported a significant dose-dependent relationship between arsenic exposure and prostate cancer i.e. RR of prostate cancer were 1.23 (95% CI, 1.16–1.30) in the medium arsenic level groups, and 1.28 (95% CI, 1.21–1.35) in the high arsenic level groups (102).

Another study evaluated breast cancer risk and drinking water contaminated by wastewater reported that individuals using a water supply of public wells had a higher breast cancer risk, however, this association was small and statistically unstable (OR = 1.4; 95% CI 0.8–2.4) (103).

Our findings, however, contradicted these findings and revealed that individuals who self-reported their source of water supply as being deep-well water wells (OR = 0.88; CI 0.60-1.31) and/or shallow well water (OR=0.29; CI 0.29-0.82) were associated with a lower risk of developing HRCs. It is important to note that the association between the consumption of shallow well water and the odds of developing HRCs was significant. From the final multivariable model

at baseline, our results showed that the association between household source of water supply and the risk of HRCs was that of borderline significance.

A family history of cancer was also explored in multivariable modeling. Our analysis showed that having a father with a positive history of cancer was associated with a higher prevalence of (OR=1.37; 95% CI = 1.01, 1.86) HRCs in the baseline study. Our data revealed a similar association between having a sibling (i.e. brother or sister) with a previous history of cancer and HRCs risk (OR = 1.51, 95% CI: 1.11, 2.07).

Our study findings are consistent with an epidemiological study examining association of family history and prostate cancer risk in a population-based cohort of Iowa men which stated that; a history of prostate cancer in father or brother was positively associated with a prostate cancer risk (RR = 3.2; 95% CI = 1.8-5.7) (104). Brewer and colleagues had found a similar relationship between breast cancer risk and a family history of cancer with a 3.5-fold increase (95% CI 2.56–4.79) risk of breast cancer (105).

As indicated earlier in this thesis that, age plays a significant role in breast cancer development (47) as well as in causing prostate cancer (63). Our study findings although observed a significant dose-response relation between age and HRCs risk, yet it should be interpreted with caution as confidence intervals appear to be very wide.

Our finding is supported by a study that reported that with advancing age there was an increased risk of carcinoma breast which doubles about every ten years until menopause (47). Contrary to this view, prostate cancer development is typically reported in older men and about 6 of every 10 cases of diagnosed prostate cancer is reported in men aged 65 or older (64). These findings indicate that age remains a risk factor when HRCs are assessed against it.

It is imperative to note in the current study that, exposure to radiation was significantly associated with more than three times the odds of developing HRCs (OR = 3.39; 95% CI: 2.23, 4.84). Our study finding was supported by Esther M. John and colleagues who reported an increased risk (OR = 3.55, CI = 1.47–8.54) of breast cancer in women who had history of radiation exposure (i.e. radiotherapy for a previous cancer) (106).

Besides, John et al confirmed a dose-response relation between breast cancer risk and exposure to radiation (107). For prostate cancer, the analogous trend was observed by one of the case-control studies examining exposure to radiation and the risk of prostate cancer (108). Their study found

that exposure to radiations (adjusted OR 2.23, 95% CI 1.42–3.49) was significantly associated with increased prostate cancer (108).

5.3 Incidence

Our study reported cumulative incidence for breast and prostate cancers in rural Saskatchewan as 0.86% and 1.08%, respectively. When combined (i.e. breast and prostate cancers together) the cumulative incidence of HRCs was 0.97%. Our finding on incidence analysis (unadjusted) was not comparable with other studies, this is because most of the reports, as discussed below, showed a relatively higher rate for HRCs.

In the 1990s, Saskatchewan recorded breast cancer cases (i.e.14,984 cases of invasive breast cancer) that were registered by the Saskatchewan Cancer Foundation during a 59-year study period (1932 to 1990) which showed a significant increase in trends over time in the overall crude (114). According to the recent cancer statistics from Saskatchewan, carcinoma breasts continued to be the most prevalent cancer-types in Saskatchewan women (29).

Similarly, for prostate cancer, an older study (1970 to 1997 inclusive) reported cancer of prostate diagnosed in the province of Saskatchewan alone (identified in the Saskatchewan Cancer Registry) showed that in 1970, the age-adjusted incidence for carcinoma prostate was 60.5 per 10^5, this rose to 101.5 per 10^5 in 1989 (116). Also, the prostate cancer incident rate steeped up in 1993, expressing a peak of 163.1 per 10^5 (116).

This trend is also true now-a-days where Saskatchewan cancer statistics at glance in the Canadian Cancer Statistics 2019 report, stated that, for men in Saskatchewan, prostate cancer is the most frequently diagnosed type of cancer, and for Saskatchewan women, breast cancer is the most frequently diagnosed type of cancer (57). According to the same report, an estimated 710 per 545,785 men will be diagnosed with prostate cancer and an estimated 730 per 552,565 women will be diagnosed with breast cancer, in the province of Saskatchewan alone (57).

In Canada, the number of new cancer cases diagnosed each year continues to increase (109). In particular, for HRCs (i.e. breast and prostate cancers), one of the reports on cancer incidence counts and rates by cancer type and sex, 2013, expressed incident rates for breast cancer in females as 13.3% with ASIR 95% CI (138.8–144.3). The same report showed the percentage of new cases of prostate cancer as 9.9% with ASIR 95% CI (115.8–121.1) (110).

Moreover, projected new cases (2019 estimates) for Canadian women identified that 128 per 10^5 women will be diagnosed with breast cancer representing 25% of all new cancer cases in women (113). This estimation of breast cancer cases reflects an average of 74 Canadian women to be diagnosed with breast cancer every day (113). Besides, for Canadian men, 118.1 per 10^5 men will be diagnosed with prostate cancer representing 20% of all new cancer cases in men and an average of 63 Canadian men will be diagnosed with prostate cancer every day (112).

In surmising findings on incident rates of breast and prostate cancers, it is refreshing to note that our data analysis revealed finding on rural Saskatchewan only and a total of 8261 participants, out of this number 247 participants at the baseline, reported as HCR cases (i.e. 117 prostate and 130 breast cancer cases), however, at the follow-up, HRC cases were 167 (i.e. 94 breast and 73 prostate cancers cases). The incidence analysis contained only new cancer cases over the four years period (i.e. 2010 to 2014) and expressed the cumulative incidence for HRCs as 0.97%.

5.4 Strengths and Limitations of the study

5.4.1 Strengths

This study established the prevalence and incidence of HRCs (i.e. breast and prostate cancers) alongside examined the potential risk factors associated with their prevalence in a rural population. The results of our study identified that work-related exposure to radiation was strongly and significantly associated with an increased risk of being diagnosed with HRCs. It provided the odds ratio to reflect the observed association between work-related exposure to radiation and the risk of HRCs development.

Our study further reported on having a sibling with a positive history of cancer was significantly associated with an elevated risk of being diagnosed with HRCs. A significant dose-response relation was observed between age and HRCs risk. The majority of our findings were comparable with the already published studies in the literature. Our study laid a foundation on which newer studies can be built upon by extending different aspects of this work.

The current study employed data from SRHS which involved the extensive collection of information on various factors i.e. individual, and contextual factors as well as the covariates. Most importantly, the SRHS was able to collect detailed information on family history of cancers. A self-administered mailed out questionnaire was used to gather the information that has been regarded by many authors as the best option for gathering data in far off areas (117).

Moreover, the content of the baseline questionnaire was optimized using a pilot's survey (88). Response rates for both the baseline and follow-up surveys were maximized using a modified version of Dillman's method for mail and telephone surveys (91). Quantitative data was mainly collected and analyzed using appropriate standard statistical techniques. The study used a concept of Health Canada's Population Health Framework (PHF) which has been successfully implemented in one of the similar cohort studies (68).

In addition, the population of SRHS involved all four quadrants of the southern half of the rural Saskatchewan. This expressed the best possible coverage of the geographical areas of the province. One of the advantages of this vast coverage is that it lessened the likelihood of bias that might have been introduced if the study was limited to one or a maximum of two quadrants of Saskatchewan.

The SRHS team was composed of researchers from a diverse background including project managers, epidemiologists, biostatisticians, nurses, geographers, the rural municipality and small town's members, all of those worked in tandem until the successful completion of the work. The heterogeneous composition of the SRHS research team extended a unique insight into this project and contributed to invaluable feedback based on one health paradigm.

5.4.2 Limitations

Like any other epidemiological study, the current study was also not void of certain limitations. The major limitation of this study was a relatively smaller number of HRC cases both in the baseline (247/8261) as well as at the follow-up (167/4867). The number of HRC cases that were identified from the SRHS was relatively small. Resultantly, low statistical power led to a compromised power of the models and consequently, our data was not able to predict time-to-HRCs (i.e. risk factors associated with the incidence of HRCs). Small statistical power could also be a reason why we obtained many statistically not-significant associations in our study.

Another limitation of the current study was that the prevalence and incidence rates were established based on a self-reported diagnosis of HRC as we were not able to link our data to the Cancer Registry. Moreover, information was available only at the two discrete points or a short follow up of four years. Dropouts from the initial cohort along with missing information were other limitations of this study.

In terms of the response rate, although it was moderate (i.e. 42%) still this would have been worked upon to better strengthen the statistical power which could facilitate more refined scrutiny of the risk factors that play role in the development of HRCs.

Occupational exposures (e.g. workplace exposure to radiations etc.) might be associated with recall bias and this could result in underestimation or overestimation of the exposure history. In addition, an overestimation of HRC cases could also occur if disease-free individuals were less likely to return completed questionnaires.

For work-related exposure to radiation, data was not collected on the specific type of radiation exposure. Meaning by, it referred to all radiation types at workplace (i.e. solar related or machinery related exposure to radiation, etc.). Since we lacked details on the specific type of radiation exposure, therefore this has been identified as one of the limitations of the current study.

The findings of our study can certainly be generalized while conducting a comparison to other parts of rural Canada where population characteristics are either alike or at least closer to Saskatchewan. At the same time, we are limited in the generalizability of our study findings to populations of urban Canada, given the fact that the population of SRHS was composed of rural inhabitants. Also, a majority of the study population was Caucasians, which poses a greater challenge when contrast is made with the rest of Canada where ethnicity is predominantly diverse.

5.5 Recommendations

As we were able to capture relatively low prevalence and incidence of HRCs in rural Saskatchewan, we would consider a case-control study design in the future to carry out the analysis. This is because of multiple reasons; a) In studies with case-control design, we can maintain a good statistical power by increasing the number of controls i.e. recruiting multiple controls for each case (provided we are limited in the number of cases obtained and data is obtainable at no added cost) (118) b) case-control study is useful to examine outcomes with long latency periods between exposure and disease manifestation (118) c) case-control study design is typically advantageous when investigation of multiple exposures is required towards the same outcome (118).

Analysis in the current project was limited to unadjusted incidence analysis only, this invites future research to find adjusted incidence for HRCs, using the same data. Besides, it would be prudent to conduct research that collects data specifically tailored to examine the potential risk

factors that play a role in the development of HRCs. This will help to predict the risk factors that could potentially predict the incidence of HRCs (i.e. time-to-HRCs).

5.6 Conclusions

The crude prevalence of HRC cases was 3.0% at the baseline and 3.4% at the follow-up. HRCs were slightly more prevalent amongst non-farm residents 3.2% when compared with farm residents 2.8%. In addition, it appeared that HRCs were more prevalent within the eastern part of the province (6.5%, 6.6%, respectively) as compared to the western part (4.6%, 6.3%, respectively), among both farm and non-farm residents, at the baseline.

In summarizing our findings, the prevalence of HRCs in this cohort was comparable with other studies conducted on rural Saskatchewan. In particular, our analysis revealed the prevalence of prostate cancer as 3.0% and 3.3% for farm & non-farm male residents at the follow-up, which was comparable with the other study conducted on rural Saskatchewan by Sharma and colleagues expressing it as 3.32% (115).

For breast cancers, the prevalence obtained in the current study, was 3.0% and 4.2% for farm & non-farm female residents, at the follow-up. Since no other study was found investigating the prevalence of breast cancers among female residents of rural Saskatchewan, therefore, this comparison was not possible in the context of rural studies of the same province.

Our study expressed cumulative incidence for breast and prostate cancers in rural Saskatchewan as 0.86% and 1.08%, respectively. The cumulative incidence for HRCs was 0.97%. In carrying out comparison for the cumulative incidence with the other studies, it was revealed that, a relatively higher rate for the cumulative incidence of HRCs have been reported.

In terms of risk factors associated with the prevalence of HRCs, our study found that; exposure to radiation and history of cancer in a first-degree relative was significantly associated with HRCs. It appeared that the prevalence of HRCs among farm and non-farm rural residents depends on the complex interplay among a variety of factors such as individual and contextual factors.

REFERENCES

1. Internet search. (2019). Available via https://www.mayoclinic.org/diseases-conditions/cancer/symptoms-causes/syc-20370588 Assessed on 3 May 2019.
2. Internet search. (2019). Available via https://www.medicinenet.com/cancer/article.htm#cancer_facts Assessed on 3 May 2019.
3. Brian E. Henderson, Heather Spencer Feigelson, Hormonal carcinogenesis, *Carcinogenesis*, Volume 21, Issue 3, March 2000, Pages 427–433, https://doi.org/10.1093/carcin/21.3.427
4. GLOBAL BURDEN OF CANCER IN WOMEN Current status, trends, and interventions Available via https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/global-cancer-facts-and-figures/global-burden-of-cancer-in-women.pdf
5. World cancer research fund – American institute for cancer research. Continuous update project (CUP) Analysis research on cancer prevention and survival. Internet search. (2019). Available via https://www.wcrf.org/dietandcancer/cancer-trends/prostate-cancer-statistics. Assessed on 3 May 2019.
6. FIRST CALIFORNIA PHYSICIAN PARTNERS (FCPP) – The top 10 cancers of America. Internet search. (2019). Available via https://fcppcentralvalley.com/cancer-center/the-top-10-cancers-of-america/. Assesses on 3 May 2019.
7. Sun, Y. S., Zhao, Z., Yang, Z. N., Xu, F., Lu, H. J., Zhu, Z. Y., … Zhu, H. P. (2017). Risk Factors and Preventions of Breast Cancer. *International journal of biological sciences*, *13*(11), 1387–1397. doi:10.7150/ijbs.21635
8. Singletary S. E. (2003). Rating the risk factors for breast cancer. *Annals of surgery*, *237*(4), 474–482. doi:10.1097/01.SLA.0000059969.64262.87
9. Leitzmann, M. F., & Rohrmann, S. (2012). Risk factors for the onset of prostatic cancer: age, location, and behavioral correlates. *Clinical epidemiology*, *4*, 1–11. doi:10.2147/CLEP.S16747
10. American Cancer Society. Internet search 2019. Available via: https://www.cancer.org/cancer/prostate-cancer/causes-risks-prevention/risk-factors.html

11. Francine Laden, Susan E. Hankinson, Donna Spiegelman, Lucas M. Neas, Graham A. Colditz, David J. Hunter, JoAnn E. Manson, Celia Byrne, Bernard A. Rosner, Frank E. Speizer, Geographic Variation in Breast Cancer Incidence Rates in a Cohort of U.S. Women, *JNCI: Journal of the National Cancer Institute*, Volume 89, Issue 18, 17 September 1997, Pages 1373–1378, https://doi.org/10.1093/jnci/89.18.1373

12. Lundqvist, A., Andersson, E., Ahlberg, I., Nilbert, M., & Gerdtham, U. (2016). Socioeconomic inequalities in breast cancer incidence and mortality in Europe-a systematic review and meta-analysis. *European journal of public health*, *26*(5), 804–813. doi:10.1093/eurpub/ckw070

13. Cheng, I., Witte, J. S., McClure, L. A., Shema, S. J., Cockburn, M. G., John, E. M., & Clarke, C. A. (2009). Socioeconomic status and prostate cancer incidence and mortality rates among the diverse population of California. *Cancer causes & control : CCC*, *20*(8), 1431–1440. doi:10.1007/s10552-009-9369-0

14. Cheng, I., Witte, J. S., McClure, L. A., Shema, S. J., Cockburn, M. G., John, E. M., & Clarke, C. A. (2009). Socioeconomic status and prostate cancer incidence and mortality rates among the diverse population of California. *Cancer causes & control : CCC*, *20*(8), 1431–1440. doi:10.1007/s10552-009-9369-0

15. Marshall J. R. (2012). Diet and prostate cancer prevention. *World journal of urology*, *30*(2), 157–165. doi:10.1007/s00345-011-0810-0

16. Internet search July_2019. Available via https://www.cancer.net/navigating-cancer-care/prevention-and-healthy-living/physical-activity-and-cancer-risk

17. Giovannucci EL, Liu Y, Leitzmann MF, Stampfer MJ, Willett WC. A Prospective Study of Physical Activity and Incident and Fatal Prostate Cancer. *Arch Intern Med.* 2005;165(9):1005–1010. doi:10.1001/archinte.165.9.1005

18. Internet search 2019 july. Available via: https://bcaction.org/our-take-on-breast-cancer/environment/

19. Breast cancer organization. Available via https://www.breastcancer.org/symptoms/understand_bc/risk/understanding

20. Kessler, Larry G. "The Relationship between Age and Incidence of Breast Cancer Population and Screening Program Data." Cancer 69.S7 (1992): 1896-903. Web.

21. Surveillance, Epidemiology, and End Results Program (SEER) of the National Cancer Institute. Fast Stats: An interactive tool for access to SEER cancer statistics. Bethesda, MD: SEER, National Cancer Institute; [Accessed November 11, 2011]. Available from: www.seer.cancer.gov/faststats. [Google Scholar]

22. Kotepui M. (2016). Diet and risk of breast cancer. *Contemporary oncology (Poznan, Poland)*, *20*(1), 13–19. doi:10.5114/wo.2014.40560

23. Marshall J. R. (2012). Diet and prostate cancer prevention. *World journal of urology*, *30*(2), 157–165. doi:10.1007/s00345-011-0810-0

24. Health and Place in Rural Canada by Allison M. Williams and Judith C. Kulig. Internet search. (2019). Available via https://www.ubcpress.ca/health-in-rural-canada. Assessed on 4 May 2019.

25. Statistics Canada. Population, urban and rural, by province and territory - Canada, census of population 1851- 2006. 2009; Available at: http://www.statcan.gc.ca/tablestableaux/sumsom/l01/cst01/demo62a-eng.htm.

26. Statistics Canada. Population, urban and rural, by province and territory- Saskatchewan, census of population 1851-2006. 2009; Available at: http://www.statcan.gc.ca/tablestableaux/sumsom/l01/cst01/demo62i-eng.htm

27. Rushton L, Hutchings SJ, Fortunato L, et al. 2012. Occupational cancer burden in Great Britain. Br J Cancer. 107(Suppl 1):S3-7.

28. Statistics Canada. Internet search. (2019). Available via https://www150.statcan.gc.ca/n1/pub/95-640-x/2016001/article/14807-eng.htm. Assessed on 4 May 2019.

29. Canadian cancer society report 2017. Canadian cancer statistics. Saskatchewan cancer statistics at glance. Internet search. (2019). Available via https://www.cancer.ca/en/cancer-information/cancer-101/canadian-cancer-statistics-publication/?region=sk. Assessed on 4 May 2019.

30. Gail P. Risbridger, Ian D. Davis, Stephen N. Birrell, and Wayne D. Tilley. "Breast and Prostate Cancer: More Similar than Different." Nature Reviews Cancer 10.3 (2010): 205-212. Web.

31. López-Abente, G., Mispireta, S., & Pollán, M. (2014). Breast and prostate cancer: an analysis of common epidemiological features in mortality trends in Spain. *BMC cancer*, *14*, 874. doi:10.1186/1471-2407-14-874

32. Dietary fat and cancer: consistency of the epidemiologic data, and disease prevention that may follow from a practical reduction in fat consumption. *Prentice RL, Sheppard L Cancer Causes Control. 1990 Jul; 1(1):81-97; discussion 99-109.*

33. Internet search 2019. Available via https://en.wikipedia.org/wiki/Carcinogenesis. Assessed on 21 July 2019.

34. MedicineNet. Internet search 2019. Available via . https://www.medicinenet.com/script/main/art.asp?articlekey=4636. Assessed on 18 July 2019.

35. NationalCancerInstitute. Internet search 2019. Available via https://www.cancer.gov/publications/dictionaries/cancer-terms/def/oncogene. Assessed on 18 July 2019.

36. Pitot, H. C. (1993). The molecular biology of carcinogenesis. *Cancer*, *72*(S3), 962-970.

37. Conference series.com. Internet search 2019. Available via https://europe.endocrineconferences.com/events-list/hormone-dependent-cancers. Assessed on 19 July 2019.

38. Russo, I H, and Russo, J. "Physiological Bases of Breast Cancer Prevention." European Journal of Cancer Prevention: The Official Journal of the European Cancer Prevention Organisation (ECP) 2 Suppl 3 (1993): 101-11. Web.

39. Russo, I. H., & Russo, J. (1996). Mammary gland neoplasia in long-term rodent studies. *Environmental health perspectives*, *104*(9), 938–967. doi:10.1289/ehp.96104938

40. Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. *CA: a cancer journal for clinicians*, *65*(2), 87-108.

41. Bray, Freddie, Peter McCarron, and D Maxwell Parkin. "The Changing Global Patterns of Female Breast Cancer Incidence and Mortality." Breast Cancer Research: BCR 6.6 (2004): 229-39. Web.

42. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015;65(2):87-108.

43. Taitt H. E. (2018). Global Trends and Prostate Cancer: A Review of Incidence, Detection, and Mortality as Influenced by Race, Ethnicity, and Geographic Location. *American journal of men's health*, *12*(6), 1807–1823. doi:10.1177/1557988318798279

44. Center MM, Jemal A, Lortet-Tieulent J, et al. International variation in prostate cancer incidence and mortality rates. *Eur Urol.* 2012;61(6):1079-1092.

45. Internet search. 2019. Available via: https://www.urotoday.com/library-resources/advanced-prostate-cancer/111566-epidemiology-and-etiology-of-prostate-cancer.html. Assessed on 15 September 2019.

46. Bray, Freddie, Jian-Song Ren, Eric Masuyer, and Jacques Ferlay. "Global Estimates of Cancer Prevalence for 27 Sites in the Adult Population in 2008." International Journal of Cancer 132.5 (2013): 1133-145. Web.

47. McPherson, K., Steel, C., & Dixon, J. M. (2000). ABC of breast diseases: breast cancer—epidemiology, risk factors, and genetics. *BMJ: British Medical Journal*, *321*(7261), 624.

48. Hilal L., Shahait M., Mukherji D., Charafeddine M., Farhat Z., Temraz S. Prostate cancer in the Arab world: A view from the inside. Clin Genitourin Cancer. 2015

49. Grönberg H. Prostate cancer epidemiology. Lancet. 2003;361:859–864

50. Parkin D.M., Muir C.S., Whelan S., Gao Y., Ferlay J., Powell J., editors. Cancer incidence in five continents, volume VI. International Agency for Research on Cancer Scientific Publications; Lyon: 1992.

51. Candian Cancer Society. Canadian Cancer Statistics A 2018 special report on cancer incidence by stage. 2018;1–52.

52. Marrett LD, De P, Airia P, Dryer D. Cancer in Canada in 2008. Cmaj. 2008;179(11):1163–70.

53. Kondro, W. (2012). Cancer incidence rises while mortality declines: CMAJ CMAJ.*Canadian Medical Association.Journal, 184*(9), E488-E489. Retrieved from http://cyber.usask.ca/login?url=https://search.proquest.com/docview/1020908529?accountid=14739

54. Noel S. Weiss, Re: "Twenty Five Year Follow-up for Breast Cancer Incidence and Mortality of the Canadian National Breast Screening Study: Randomised Screening Trial", *American Journal of Epidemiology*, Volume 180, Issue 7, 1 October 2014, Pages 759–760, https://doi-org.cyber.usask.ca/10.1093/aje/kwu227

55. Ellison, L. F., & Wilkins, K. (2009). Cancer prevalence in the Canadian population. *Health Reports*, *20*(1), 7

56. Canadian Cancer Society, Canadian Cancer Society's Advisory Committee on Cancer Statistics. Canadian Cancer Statistics 2017. Can Cancer Soc [Internet]. 2017;2017:1–132. Available from: http://www.cancer.ca/~/media/cancer.ca/CW/publications/Canadian Cancer Statistics/Canadian-Cancer-Statistics-2017-EN.pdf%0Acancer.ca/Canadian-CancerStatistics-2017-EN.pdf

57. Canadian Cancer Statistics Advisory Committee. *Canadian Cancer Statistics 2019.* Toronto, ON: Canadian Cancer Society; 2019. Available at: cancer.ca/Canadian-Cancer-Statistics-2019-EN. Assessed on 15 September 2019.

58. Internet search. 2019. Available via https://www.merriam webster.com/dictionary/etiologic. Assessed on 8, July 2019.

59. Internet search. 2019. Available via https://www.who.int/topics/risk_factors/en/. Assessed on 8, July 2019.

60. Song Wu, Scott Powers, Wei Zhu, and Yusuf A. Hannun. "Substantial Contribution of Extrinsic Risk Factors to Cancer Development." Nature 529.7584 (2015): 43-47J. Web.

61. Medicine Net. Cancer risk factors. Internet search. (2019). Available via https://www.medicinenet.com/cancer_causes/article.htm. Assessed on 7 June 2019.

62. Bostwick, D. G., Burke, H. B., Djakiew, D., Euling, S., Ho, S. M., Landolph, J., ... & Timms, B. (2004). Human prostate cancer risk factors. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, *101*(S10), 2371-2490.

63. Prostate Cancer UK. Your diet and physical activity. Internet search. (2019). Available via https://prostatecanceruk.org/prostate-information/living-with-prostate-cancer/your-diet-and-physical-activity. Assessed on 9 June 2019.

64. American Cancer Society. About Prostate Cancer. Internet search. (2019). Available via https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html. Assessed on 8 May 2019.

65. Breast Cancer Organization. Internet search. (2019). Available via https://www.breastcancer.org/risk/factors/family_history. Assessed on 8 June 2019.

66. Pahwa P, Karunanayake CP, Hagel L, Janzen B, Pickett W, Rennie D, et al. The Saskatchewan rural health study: an application of a population health framework to understand respiratory health outcomes. BMC Res Notes [Internet]. 2012;5:400.

67. Scott L, Mobley LR, Il'Yasova D. Geospatial analysis of inflammatory breast cancer and associated community characteristics in the United States. Int J Environ Res Public Health. 2017;14(4):1–10.

68. Zahnd, W., James, A., Jenkins, W., Izadi, S., Fogleman, A., Steward, D., . . . Brard, L. (2018). Rural-Urban Differences in Cancer Incidence and Trends in the United States. Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology, 27(11), 1265-1274.

69. Gray, R. E., James, P., Manthorne, J., Gould, J., & Fitch, M. I. (2004). A consultation with Canadian rural women with breast cancer. *Health expectations: an international journal of public participation in health care and health policy*, *7*(1), 40–50. doi:10.1046/j.1369-6513.2003.00248.x

70. Nguyen-Pham, Leung, & Mclaughlin. (2014). Disparities in breast cancer stage at diagnosis in urban and rural adult women: A systematic review and meta-analysis. Annals of Epidemiology, 24(3), 228-235.

71. What Can Geography Tell Us About Prostate Cancer? Klassen, Ann C. et al. American Journal of Preventive Medicine, Volume 30, Issue 2, S7 - S15

72. American Cancer Society. Internet search. (2019). Available via https://www.cancer.org/cancer/prostate-cancer/causes-risks-prevention/risk-factors.html. Assessed on 8 June 2019.

73. Cheng, I., Witte, J. S., McClure, L. A., Shema, S. J., Cockburn, M. G., John, E. M., & Clarke, C. A. (2009). Socioeconomic status and prostate cancer incidence and mortality rates among the diverse population of California. *Cancer causes & control: CCC*, *20*(8), 1431–1440. doi:10.1007/s10552-009-9369-0

74. Lundqvist, A., Andersson, E., Ahlberg, I., Nilbert, M., & Gerdtham, U. (2016). Socioeconomic inequalities in breast cancer incidence and mortality in Europe-a systematic review and meta-analysis. *European journal of public health*, *26*(5), 804–813. doi:10.1093/eurpub/ckw070.

75. American Cancer Society. Diet and Physical Activity. Whats the cancer connection? Internet search. (2019). Available via https://www.cancer.org/cancer/cancer-causes/diet-physical-activity/diet-and-physical-activity.html. Assessed on 9 June 2019.

76. Canadian Medical Association policy 2015. Internet search. 2019. Available via: https://www.cma.ca/sites/default/files/2018-11/PD15-12.pdf . Assessed on 15 September 2019.

77. Fournier A, Dos Santos G, Guillas G, et al. Recent recreational physical activity and breast cancer risk in postmenopausal women in the E3N cohort. *Cancer Epidemiology, Biomarkers & Prevention* 2014; 23(9):1893-1902. [PubMed Abstract]

78. Wu Y, Zhang D, Kang S. Physical activity and risk of breast cancer: a meta-analysis of prospective studies. *Breast Cancer Research and Treatment* 2013; 137(3):869-882.

79. Gann PH, Hennekens CH, Sacks FM, Grodstein F, Giovannucci E, Stampfer MJ. Prospective study of plasma fatty acids and risk of prostate cancer. J Natl Cancer Inst. 1994;86:281–286.

80. Marshall J. R. (2012). Diet and prostate cancer prevention. *World journal of urology*, *30*(2), 157–165. doi:10.1007/s00345-011-0810-0

81. Kotepui M. (2016). Diet and risk of breast cancer. *Contemporary oncology (Poznan, Poland)*, *20*(1), 13–19. doi:10.5114/wo.2014.40560

82. Pukkala E, Martinsen JI, Lynge E, Gunnarsdottir HK, Sparn P, Tryggvadottir L, et al. Occupation and cancer follow-up of 15 million people in five Nordic countries. Vol. 48, Acta Oncologica. 2009. 646-790 p.

83. H Shimizu, Rk Ross, L Bernstein, R Yatani, Be Henderson, & Tm Mack. (1991). Cancers of the prostate and breast among Japanese and white immigrants in Los Angeles County. British Journal of Cancer, 63(6), 963-966.

84. Meyer TE, Coker AL, Sanderson M, Symanski E. A case-control study of farming and prostate cancer in African-American and Caucasian men. Occup Environ Med. 2007;64(3):155–60.

85. Blair A, Zahm SH, Pearce NE, Heineman EF, Fraumeni JF. Clues to cancer etiology from studies of farmers. Scand J Work Environ Heal. 1992;18(4):209–15.

86. Ritter L, Wigle Dt, Semenciw Rm, Wilkins K, and Mao Y. "Mortality Study of Canadian Male Farm Operators - Cancer Mortality and Agricultural Practices in Saskatchewan." (1990): 499-505. Web.

87. Pahwa, P., Karunanayake, C., Hagel, L., Janzen, B., Pickett, W., Rennie, D., . . . Dosman, J. (2012). The Saskatchewan rural health study: An application of a population health framework to understand respiratory health outcomes. BMC Research Notes, 5(1), 400.

88. Pahwa, P., Karunanayake, C., Hagel, L., Janzen, B., Rennie, D., Lawson, J., . . . Dosman, J. (2012). Self-Selection Bias in an Epidemiological Study of Respiratory Health of a Rural Population. Journal of Agromedicine, 17(3), 316-325.

89. du Plessis V, Beshiri R, Bollman RD, Clemenson H: Definitions of "Rural". Agriculture and Rural Working Paper Series, Working Paper No. 61. Catalogue no. 21-601-MIE- No. 061. Ottawa, Canada: Agricultural Division, Statistics Canada; 2004.

90. Pahwa, Punam, Rana, Masud, Pickett, William, Karunanayake, Chandima P., Amin, Khalid, Rennie, Donna, . . . Dosman, James. (2017). Cohort profile: The Saskatchewan Rural Health Study--adult component. BMC Research Notes, 10(1), 1-7.

91. Dillman, D. (2000). Mail and internet surveys: The tailored design method (2nd ed.). New York: J. Wiley.

92. Janout, V., & Kollárová, H. (2001). Epidemiology of colorectal cancer. Biomedical Papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia, 145(1), 5-10.

93. Pahwa, P. (2000). Statistical modelling of longitudinal lung function data (Doctoral dissertation). University of Saskatchewan, Saskatoon, Canada.

94. Liang, K., & Zeger, S. (1986). Longitudinal Data Analysis Using Generalized Linear Models. Biometrika, 73(1), 13-22.

95. Fitzmaurice GM, Laird NM, Ware JH. Applied Longitudinal Analysis. John Wiley & Son, Inc. 2002.

96. Rothman, K.J. and S. Greenland, Modern Epidemiology. 1998: Lippincott-Raven.

97. Hosmer, David W., Lemeshow, Stanley, & Sturdivant, Rodney X. (2013). Model-Building Strategies and Methods for Logistic Regression. In Wiley Series in Probability and Statistics (pp. 89-151). Hoboken, NJ, USA: John Wiley & Sons.

98. Kleinbaum, D., Klein, Mitchel. author, & SpringerLink. (2005). Survival Analysis: A Self-Learning Text (Second ed., Statistics for Biology and Health).

99. Cox, D. (1972). Regression Models and Life-Tables. Journal of the Royal Statistical Society. Series B (Methodological), 34(2), 187-220.

100. Leon Gordis. (2009). Gordis: Epidemiology (4th edition). Saunders Elsevier. ISBN:978-1-4160-4002-6

101. Cantor, K. (1997). Drinking water and cancer. Cancer Causes & Control, 8(3), 292-308.

102. Roh, T., Lynch, C., Weyer, P., Wang, K., Kelly, K., & Ludewig, G. (2017). Low-level arsenic exposure from drinking water is associated with prostate cancer in Iowa. Environmental Research, 159, 338-343.

103. Brody, J., Aschengrau, A., McKelvey, W., Swartz, C., Kennedy, T., & Rudel, R. (2006). Breast cancer risk and drinking water contaminated by wastewater: A case control study. Environmental Health: A Global Access Science Source, 5, 28.

104. Cerhan, J., Parker, A., Putnam, S., Chiu, B., Lynch, C., Cohen, M., . . . Cantor, K. (1999). Family history and prostate cancer risk in a population-based cohort of Iowa men. Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology, 8(1), 53-60.

105. Brewer, H., Jones, R., Schoemaker, M., Ashworth, E., & Swerdlow, M. (2017). Family history and risk of breast cancer: An analysis accounting for family structure. Breast Cancer Research and Treatment, 165(1), 193-200

106. John, E., Phipps, A., Knight, J., Milne, R., Dite, G., Hopper, J., . . . Whittemore, A. (2007). Medical radiation exposure and breast cancer risk: Findings from the Breast Cancer Family Registry. International Journal of Cancer, 121(2), 386-394.

107. Boice, J., Land, C., Shore, R., Norman, J., & Tokunaga, M. (1979). Risk of breast cancer following low-dose radiation exposure. Radiology, 131(3), 589-597.

108. P Myles, S Evans, A Lophatananon, P Dimitropoulou, D Easton, T Key, . . . K Muir. (2008). Diagnostic radiation procedures and risk of prostate cancer. British Journal of Cancer, 98(11), 1852-6.

109. Internet search. (2019). Available via https://www.canada.ca/content/dam/phac-aspc/documents/services/publications/diseases-conditions/fact-sheet-cancer-canada/fact-sheet-cancer-canada.pdf. Assessed on 6 November 2019.

110. Cancer Care Ontario. Ontario Cancer Statistics 2016. Toronto: Cancer Care Ontario; 2018. ISSN 2371-039X Key title: Ontario cancer statistics (Print) ISSN 2371-0403 Key title: Ontario cancer statistics (Online) (2019). Available via https://www.cancercareontario.ca/sites/ccocancercare/files/assets/OCSChapterFive.pdf

111. Internet search. (2019). Available via https://nationalpost.com/features/cancer-in-canada. Assessed in 6 November 2019.

112. Internet search. (2019). Available via https://www.cancer.ca/en/cancer-information/cancer-type/prostate/statistics/?region=sk. Assessed on 6 November 2019.

113. Internet search. (2019). Available via http://www.cancer.ca/en/cancer-information/cancer-type/breast/statistics/?region=sk. Assessed on 6 November 2019.

114. Wang, P., & Cao, P. (1996). Incidence trends of female breast cancer in Saskatchewan, 1932–1990. Breast Cancer Research and Treatment, 37(3), 197-207.

115. Sharma, M., Lawson, J., Kanthan, R., Karunanayake, C., Hagel, L., Rennie, D., . . . Pahwa, P. (2016). Factors Associated With the Prevalence of Prostate Cancer in Rural Saskatchewan: The Saskatchewan Rural Health Study. Journal of Rural Health, 32(2), 125-135.

116. Skarsgard, D., & Tonita, J. (2000). Prostate cancer in Saskatchewan Canada, before and during the PSA era. Cancer Causes & Control, 11(1), 79-88.

117. Dillman, DA. (1978). Mail and telephone surveys: The total design method. New York: Wiley.

118. Setia M. S. (2016). Methodology Series Module 2: Case-control Studies. *Indian journal of dermatology*, *61*(2), 146–151. doi:10.4103/0019-5154.177773

119. Fontenele, E.G.P.; Martins, M.R.A.; Quidute, A.R.P.; Montenegro, R., Jr. Contaminantes ambientais e os interferentes endócrinos. Arq. Bras. Endocrinol. Metab. **2010**, 54, 6–16. [CrossRef] [PubMed]

120. Abubakari, I., & University of Saskatchewan, College of Graduate Studies Research. (2019). Prevalence and Incidence Of Colorectal Cancer In Rural Saskatchewan: An Application Of Generalized Estimating Equations (Gee) And Survival Analysis.

121. Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13-22.

122. Canadian cancer society estimates 2019. Canadian cancer statistics. Internet search. (2020). Available via: https://www.cancer.ca/en/cancer-information/cancer-101/cancer-statistics-at-a-glance/?region=sk. Assessed on 29 February 2020.

123. Tian, J. M., Ran, B., Zhang, C. L., Yan, D. M., & Li, X. H. (2018). Estrogen and progesterone promote breast cancer cell proliferation by inducing cyclin G1 expression. *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas*, *51*(3), 1–7. https://doi.org/10.1590/1414-431X20175612

124. Gordon, E. M., Ravicz, J. R., Liu, S., Chawla, S. P., & Hall, F. L. (2018). Cell cycle checkpoint control: The cyclin G1/Mdm2/p53 axis emerges as a strategic target for broad-spectrum cancer gene therapy - A review of molecular mechanisms for oncologists. *Molecular and clinical oncology*, *9*(2), 115–134. https://doi.org/10.3892/mco.2018.1657.

125. Internet search. (2020). Available via: https://www.healthywomen.org/condition/androgen. Assessed on 1 March 2020.

126. Shi, Y., Han, J. J., Tennakoon, J. B., Mehta, F. F., Merchant, F. A., Burns, A. R., Howe, M. K., McDonnell, D. P., & Frigo, D. E. (2013). Androgens promote prostate cancer cell growth through induction of autophagy. *Molecular endocrinology (Baltimore, Md.)*, *27*(2), 280–295. https://doi.org/10.1210/me.2012-1260

127. Net Doctor. How female hormones change over time. Internet search. (2020). Available via: https://www.netdoctor.co.uk/healthy-living/a11666/female-hormones/. Assessed on 1 March 2020.

128. Wang, Y., Dai, B., & Ye, D. (2015). CHEK2 mutation and risk of prostate cancer: a systematic review and meta-analysis. *International journal of clinical and experimental medicine*, *8*(9), 15708–15715.

Appendix A: The SRHS Baseline Questionnaire

SASKATCHEWAN RURAL HEALTH STUDY

UNIVERSITY OF SASKATCHEWAN

TO MEMBERS OF THE HOUSEHOLD AND THEIR FAMILIES:

The University of Saskatchewan is conducting this project to learn more about the health of rural dwellers in Saskatchewan. Families from across Saskatchewan are participating.

This questionnaire is our first contact with your family. Please have an adult family member complete this part of the questionnaire. Please try to answer all of the questions, but remember you don't have to answer any questions if you choose not to. When you have finished, place the questionnaire in the enclosed stamped envelope and mail it back to us at the University.

Instructions

1. Please have an adult family member (age 18 or over) complete Section A and Section B of this questionnaire.

   In Section B of this form, we have asked questions about each adult member (age 18 or over) of your family. We have included enough space in this booklet for 2 adults.

   If you have more than 2 adult family members living in your home, PLEASE COMPLETE "Section B" IN THE GREEN BOOKLET for each additional adult.

2. Please read each question carefully.

3. Answer each question by placing a check mark in the box provided. For some questions you will need to write in the space provided. Thank you for taking part in this important study.

4. **Please be sure to complete the last page.**

The University of Saskatchewan

Sponsored by the Canadian Institutes of Health Research
(Canada's main funder of medical research)

SECTION A YOUR HOME

PLEASE ANSWER THE FOLLOWING QUESTIONS ABOUT YOUR PRIMARY FAMILY HOME - THAT IS THE HOME WHERE YOU LIVE MOST OF THE TIME.

Today's Date: _____/_____/_____
(Day / Month / Year)

DEMOGRAPHICS

A-1 Where is your home located?
- c Farm
- c In town
- c Acreage, please specify number of acres _____

A-2 How many people live in your home?
_____ Number

A-3 Please list all persons who usually live here including yourself.

Age	Sex	Family Member
	M c F c	Yes c No c
	M c F c	Yes c No c
	M c F c	Yes c No c
	M c F c	Yes c No c
	M c F c	Yes c No c
	M c F c	Yes c No c

(*IF MORE SPACES ARE REQUIRED CONTINUE ON THE BACK OF THE QUESTIONNAIRE.*)

A-4 How many bedrooms do you have in your home?
_____ Number

A-5 Do you own your home?
- c Yes
- c No
- c Don't know

LIVING ENVIRONMENT

A-6 What year was your residence/apartment built (approximately)?

Year_____ Don't know c

A-7 What are the types of fuel sources used to heat your home? **Please check all that apply.**

	Primary	Secondary
c Natural Gas	c	c
c Propane	c	c
c Electricity	c	c
c Fuel oil	c	c
c Coal	c	c
c Geo-thermal	c	c
c Solar energy	c	c
c Wood	c	c
➥ If yes, do you use: c Fireplace c Free standing wood stove c Fireplace insert c Outdoor wood stove		
c Other	c	c
Please specify _____		
c Don't Know		

A-8 Does your heating system have a filter?
- c Yes
- c No
- c Don't Know

A-9 Does your home have air conditioning?
- c Yes → **If yes, please check one:**
  c Central c Room c Both
- c No
- c Don't Know

A-10 Is a humidifier or vaporizer used in your home?
- c Yes
- c No
- c Don't Know

A-11 Do you use a dehumidifier in your home?
- c Yes
- c No
- c Don't Know

A-12 On average, how often per month:
do you vacuum carpet? ____ times per month
do you mop smooth floors? ____ times per month
do you dry dust clean? _____ times per month do you wet dust clean? _____ times per month

A-13 During the past 12 months, has there been water or dampness in your home from broken pipes, leaks, heavy rain, or floods?
- c Yes
- c No
- c Don't Know

Your Home

A-14 Does your home (including basement) frequently have a mildew odor or musty smell?

- c Yes
- c No
- c Don't Know

A-15 In the past 12 months have you had any of the following pets living in your home? **Please check Yes or No for each type of pet.**

c Check here if you do not have any pets in the house.

	Yes	No	Don't Know
Cat	c	c	c
Dog	c	c	c
Bird	c	c	c
Any other pet	c	c	c

If Yes, please specify _____

A-16 Within the past 12 months, were pesticides (including herbicides, insecticides, fungicides, rodenticides, fumigants) applied inside your residence (e.g., raid, spider bait, ant bait, rat bait)?

- c Yes → **If Yes, what pesticide(s)? Please specify**_____
- c No
- c Don't Know

A-17 Do any of the people who live in your house use any of the following tobacco products in the home? **Please answer Yes or No for each product.**

	Yes	No	Don't Know
Cigarettes	c	c	c
Cigars	c	c	c
Pipes	c	c	c

A-18 **If yes to cigarettes,** how many persons smoke cigarettes in your home?

_____ number of persons

A-19 **If yes to cigarettes,** how many cigarettes do they smoke per day in total?

_____ number of cigarettes

A-20 What is your best estimate of the total income, before taxes and deductions, of all household members from all sources in the past 12 months?

- c Less than $14,999
- c $15,000 to $19,999
- c $20,000 to $29,999
- c $30,000 to $39,999
- c $40,000 to $49,999
- c $50,000 to $59,999
- c $60,000 to $79,999
- c $80,000 or more

A-21 At the end of the month, how much money do you have left over? **(Please check only one)**

- c Some money
- c Just enough money
- c Not enough money

ACCESS TO HEALTH CARE

A-22 Do you and your family members in your household have access to a regular family doctor or nurse practitioner?

- c Yes
- c No
- c Don't Know

A-23 In the past 12 months did you ever experience any difficulties getting the routine or on-going care you or a family member in your household needed?

- c Yes
- c No
- c Don't Know

A-24 In the past 12 months, have you required a visit to a medical specialist for a diagnosis or consultation for yourself or a family member in your household?

- c Yes
- c No → **If No, go to question A-28.**
- c Don't Know

A-25 In the past 12 months did you ever experience any difficulty getting the specialist care you needed for a diagnosis or consultation for yourself or a family member in your household?

Yes c
No c
Don't Know c

A-26 In the past 12 months, have you or a family member in your household required immediate 24 hour health care services for a medical emergency? c Yes

c No → **If No, go to question A-30.**

c Don't know

A-27 In the past 12 months, did you ever experience any difficulties getting immediate 24 hour health care services for a medical emergency for yourself c or a family member in your household?

Yes

c No

c Don't know

A-28 How far do you travel to receive routine and ongoing medical care?_____ Km

A-29 How far do you travel to receive 24 hour emergency health care services?_____ Km

A-30 How far do you travel to receive medical or surgical specialist services?_____ Km

A-31 On average, how long does it take for an ambulance to arrive at your home in an emergency? _____minutes c Don't Know

OUTDOOR ENVIRONMENT

A-32 Do you have an indoor (barn) intensive livestock operation (building) located near your home?

c Yes → **If Yes, how far?**

c Within 1/4 mile c Greater than 1/4 mile c No

c Don't know

A-33 Do you have an outdoor feedlot or corrals located near your home?

c Yes → **If Yes, how far?**

c Within 1/4 mile c Greater than 1/4 mile c No

c Don't know

A-34 Do you have a balestack or bales located near your home?

c Yes → **If Yes, how far?**

c Within 1/4 mile c Greater than 1/4 mile c No

c Don't know

A-35 Do you have grain bins located near your home?

c Yes → **If Yes, how far?**

c Within 1/4 mile c Greater than 1/4 mile

c No

c Don't know

A-36 Do you have a sewage pond or manure lagoon located near your home?

c Yes → **If Yes, how far?**

c Within 1/4 mile c Greater than 1/4 mile

c No

c Don't know

A-37 What is the **main** source of the water supply for drinking purposes in your home?

c Bottled water
c Deep well water (more than 100 ft)
c Shallow well water (less than 100 ft)
c Spring, river or creek
c Dugout, reservoir
c Lake
c Other source:
Please specify_____

PLEASE COMPLETE THIS SECTION IF YOU LIVE ON A FARM.

FARM DEMOGRAPHICS

A-38 From the list below, please check each commodity that is produced for sale on your farm or ranch **(Please check all that apply).**

c Grain crops
c Cattle (beef)
c Cattle (dairy)
c Pigs
c Poultry
c Vegetable/Fruit
c Other:
Please specify_____

A-39 What is the area of land in your operation that you farmed or ranched last growing season? **(Please exclude land rented to others).**

Grain crops	___ acres
Forage crops	___ acres
Pasture	___ acres
Summerfallow	___ acres
Other	___ acres

A-40 How many of these types of livestock are typically raised on your farm?

No livestock	c
Cattle (beef)	____ number
Cattle (dairy)	____ number
Pigs	____ number
Poultry	____ number
Other	____ number

THIS CONCLUDES SECTION A. PLEASE PROCEED TO SECTION B, ADULT 1(GREEN TAB).

SECTION B INDIVIDUAL QUESTIONS

WE WOULD LIKE TO KNOW ABOUT EACH ADULT FAMILY MEMBER (18 YEARS OR OVER) LIVING IN YOUR HOUSEHOLD. IN THIS BOOKLET, WE HAVE INCLUDED SPACE FOR 2 ADULTS.

IF YOU HAVE MORE THAN 2 ADULT FAMILY MEMBERS LIVING IN YOUR HOME, PLEASE COMPLETE "Section B" IN THE GREEN BOOKLET FOR EACH ADDITIONAL ADULT.

ADULT 1

NOW, PLEASE ANSWER THE FOLLOWING QUESTIONS ABOUT ADULT # 1.

B-1 Age as of January 1st, 2010: _____

B-2 Date of birth: MM_____ DD____ YY_____

B-3 Sex: Male c Female c

B-4 Highest level of education:

- c Less than high school
- c Completed high school
- c Completed university
- c Completed post-secondary education other than above

B-5 What is your ethnic background?

- c Caucasian
- c First Nation
- c Metis
- c Other → **Please specify:** _____

B-6 What is your height? _____cm. **OR** _____ft and in.

B-7 What is your weight? _____Kg. **OR** _____lbs

B-8 What is your marital status? **(Please check only one)**

- c Married
- c Common law/living together
- c Widowed
- c Divorced/separated
- c Single, never married

RESPIRATORY HEALTH

COUGH

B-9 Do you usually have a cough?

- c Yes
- c No → **If no, go to question B-12.**

B-10 Do you usually cough like this on most days for 3 consecutive months or more during the year?

- c Yes
- c No

B-11 For how many years have you had this cough?

_____ years

PHLEGM

B-12 Do you usually bring up phlegm from your chest? c Yes

c No → **If no, go to question B-15.**

B-13 Do you bring up phlegm like this on most days for 3 consecutive months or more during the year?

- c Yes
- c No

B-14 For how many years have you had trouble with phlegm?

_____ years

WHEEZE

B-15 Does your chest ever sound wheezy or whistling:

	Yes	No
1. When you have a cold?	c	c
2. Apart from colds?	c	c
3. Most days or nights?	c	c

If YES to 1, 2, OR 3, for how many years has this been present? _____number of years

B-16 Have you ever had an attack of wheezing that has made you feel short of breath?

- c Yes
- c No

If YES, have you ever required medicine or treatment for the(se) attack(s)?

- c Yes
- c No

BREATHLESSNESS

B-17 Are you troubled by shortness of breath when hurrying on the level or walking up a slight hill?

- c Yes
- c No

B-18 Do you have to walk slower than people of your age because of breathlessness?

- c Yes
- c No

B-19 Do you ever have to stop for breath when walking at your own pace on the level?

- c Yes
- c No

B-20 Do you ever have to stop for breath after walking about 100 yards (or after a few minutes) on the level?

c Yes

c No

B-21 Are you too breathless to leave the house or breathless on dressing or undressing?

c Yes

c No

ASTHMA

B-22 Have you ever had asthma? c Yes

c No → **If no, go to question B-26.**

B-23 **If Yes to B-22:**

Do you still have it? c Yes c No
Was it confirmed by a doctor? c Yes c No

At what age did it start? ____ age in years

If you no longer have it, at what age did it stop? ____ age in years

B-24 **If yes to B-22**, how many times have you required services for asthma from the following places during the past 12 months?

Hospital inpatient: _____ times
Emergency room outpatient: _____ times
Doctor's office: _____ times

B-25 **If yes to B-22,** which of the following statements best describes your asthma medication use in the past 12 months:

c Never in the past 12 months
c At least once in the past 12 months
c At least once per month
c At least once per week
c Every day

ALLERGIES

B-26 Have you ever had an allergic reaction to any of the following: **(Please check all that apply).**

1. House dust c Yes c No
2. Cats c Yes c No
3. Dogs c Yes c No
4. Grasses c Yes c No
5. Pollens c Yes c No
6. Molds c Yes c No
7. Others, c Yes c No

Please specify:_____

PHYSICAL ACTIVITY

B-27 Do you exercise?

c Yes → **If yes, how many times a week?** _____times a week

c No → **If no, go to question B-29.**

B-28 How long do you usually exercise?

c Less than 15 minutes
c 15 to 30 minutes
c 31 to 60 minutes
c More than 60 minutes
c Don't Know

B-29 In a **typical week** in the past **3 months**, how much time did you usually spend on a computer, including playing computer games and using the Internet or World Wide Web? **(Please do not include time spent at work or at school)**

c None
c Less than 1 hour
c From 1 to 2 hours
c From 3 to 5 hours
c From 6 to 10 hours
c From 11 to 14 hours
c From 15 to 20 hours
c More than 20 hours

B-30 In a **typical week** in the past **3 months**, how much time did you usually spend watching television or videos?

c None
c Less than 1 hour
c From 1 to 2 hours
c From 3 to 5 hours
c From 6 to 10 hours
c From 11 to 14 hours
c From 15 to 20 hours
c More than 20 hours

EARLY LIFE EXPOSURES

B-31 Have you ever lived on a farm?

c Yes

c No

c Don't know

B-32 Did you live on a farm during your first year of life?

c Yes → **If yes, what type of farm? (Check all that apply)**

c Grain
c Livestock

c No

c Don't know

B-33 Did your mother smoke while she was pregnant with you?

c Yes

c No

c Don't know

Adult 1

B-34 What was your birth weight?
_____ pounds or _____ grams

c Don't know

B-35 Were you breastfed as a child?

c Yes → **If yes, was it for 6 months or longer?** c Yes c No

c No

c Don't know

CIGARETTE SMOKING

B-36 Have you ever smoked cigarettes? **(If you have smoked less than 20 packs of cigarettes in your lifetime, answer no.)**

c Yes

c No → **If no, go to question B-43**

B-37 Do you now smoke cigarettes?

c Yes

c No

B-38 How old were you when you first started regular cigarette smoking? ____ years old

B-39 How many cigarettes do you smoke per day now?_____ cigarettes per day

B-40 On the average of the entire time you smoked, how many cigarettes did you smoke per day?
_____ cigarettes per day

B-41 If you have stopped smoking cigarettes completely, how old were you when you stopped?
_____ age stopped

B-42 If there have been periods when you abstained from smoking, indicate total years of abstinence from smoking. _____ years

B-43 Have you ever smoked a pipe regularly? **(Yes means more than 12 oz of tobacco in a lifetime)**

c Yes

c No

B-44 Have you ever smoked cigars regularly? **(Yes means more than 1 cigar a week for a year)**

c Yes

c No

B-45 Do you smoke a pipe or cigars regularly at present?

c Yes

c No

ALCOHOL CONSUMPTION

B-46 During the past 12 months, how often did you drink alcoholic beverages?

c Never
c Less than once a month
c Once a month
c 2 to 3 times a month
c Once a week
c 2 to 3 times a week
c 4 to 6 times a week
c Every day

B-47 How often in the past 12 months have you had 5 or more drinks on one occasion?

c Never
c Less than once a month
c Once a month
c 2 to 3 times a month
c Once a week
c More than once a week

MEDICAL HISTORY

B-48 In general would you say your health is:

c Excellent
c Very Good
c Good
c Fair
c Poor

B-49 During the <u>past 12 months</u>, were you seen by a doctor or other primary care giver for:

	Yes	No	Don't know
Stomach acidity or reflux?	c	c	c
An ear infection?	c	c	c
An injury?	c	c	c

B-50 Has a doctor or primary care giver ever said you have:

	Yes	No	Don't Know
Diabetes	c	c	c
Heart Disease	c	c	c
Heart Attack	c	c	c
Hardening of the arteries	c	c	c
High Blood Pressure	c	c	c
Cystic Fibrosis	c	c	c
Tuberculosis	c	c	c
Stroke	c	c	c
Cancer	c	c	c

If yes to cancer, please specify type(s):

_____

_____

CHEST ILLNESSES

B-51 Has a doctor ever said you had any of the following chest illnesses:

	Chest Illness	During the Past 12 Months		Ever In Your Life	
a.	Attack of bronchitis	c Yes	c No	c Yes	c No
b.	Pneumonia	c Yes	c No	c Yes	c No
c.	Hay Fever	c Yes	c No	c Yes	c No
d.	Sinus Trouble	c Yes	c No	c Yes	c No
e.	Chronic Bronchitis	c Yes	c No	c Yes	c No
f.	Emphysema	c Yes	c No	c Yes	c No
g.	COPD (Chronic Obstructive Pulmonary Disease)	c Yes	c No	c Yes	c No
h.	Sleep Apnea	c Yes	c No	c Yes	c No
i.	Other Chest Illness (Example chest operation) please specify:	c Yes	c No	c Yes	c No

B-52 **If yes to Chronic Obstructive Pulmonary Disease (COPD) in question B-51g,** how many times have you required services for COPD from the following places during the **past 12 months**?

Hospital inpatient: _____ times

Emergency room outpatient: _____ times

Doctor's office: _____ times

REST AND SLEEP

B-53 Do you snore?

c Yes

c No → **If no, go to question B-55.**

c Don't know

B-54 If you snore, is your snoring:

c Slightly louder than breathing?

c As loud as talking?

c Louder than talking?

c Very loud - can be heard in adjacent rooms?

B-55 How likely are you to doze off or fall asleep in the situations described below, in contrast to just feeling tired? This refers to your usual way of life in recent times. Even if you haven't done some of these things recently, try to work out how they would have affected you. **Please check one response choice for each situation.**

SITUATION	RESPONSE CHOICES			
	Would never doze	Slight chance of dozing	Moderate chance of dozing	High chance of dozing
Sitting and reading	c	c	c	c
Watching TV	c	c	c	c
Sitting inactive in a public place (e.g., a theatre or a meeting)	c	c	c	c
As a passenger in a car for an hour without a break	c	c	c	c
Lying down to rest in the afternoon when circumstances permit	c	c	c	c
Sitting and talking to someone	c	c	c	c
Sitting quietly after lunch without alcohol	c	c	c	c
In a car, while stopped for a few minutes in the traffic	c	c	c	c

Adult 1

FAMILY HISTORY

B-56 Have the following members of your biological family ever had:

	FATHER			MOTHER			BROTHER/SISTER		
	Yes	No	Don't Know	Yes	No	Don't Know	Yes	No	Don't Know
Diabetes	c	c	c	c	c	c	c	c	c
Heart Disease	c	c	c	c	c	c	c	c	c
Heart Attack	c	c	c	c	c	c	c	c	c
Hardening of the arteries	c	c	c	c	c	c	c	c	c
High Blood Pressure	c	c	c	c	c	c	c	c	c
Cystic Fibrosis	c	c	c	c	c	c	c	c	c
Tuberculosis	c	c	c	c	c	c	c	c	c
Stroke	c	c	c	c	c	c	c	c	c
Lung Trouble (Asthma,Emphysema, Chronic Bronchitis)	c	c	c	c	c	c	c	c	c
Cancer	c	c	c	c	c	c	c	c	c
If yes to cancer, please specify type(s):	_____			_____			_____		

OCCUPATIONAL HISTORY

B-57 Please list all full-time jobs at which you have worked for at least one year, starting with your present or most recent job. Please state the job title and business as specifically as possible. For example, 'mixed farming' instead of 'farming'.

Job Title	Business, Industry or Service	Total number of Years at job
e.g. Nurse	Health Care	10
e.g. Farmer	Mixed Farming	30

B-58 Have you ever been exposed to any of the following in the work place?

	No	Yes	If Yes, how often?		How many years?
Grain Dust	c	c	Daily c Weekly c	Monthly c Occasionally c	
Mine dust (e.g. potash, uranium) Specify _____	c	c	Daily c Weekly c	Monthly c Occasionally c	
Asbestos dust	c	c	Daily c Weekly c	Monthly c Occasionally c	
Wood dust	c	c	Daily c Weekly c	Monthly c Occasionally c	
Other dust Specify _____	c	c	Daily c Weekly c	Monthly c Occasionally c	
Livestock	c	c	Daily c Weekly c	Monthly c Occasionally c	
Smoke from stubble burning	c	c	Daily c Weekly c	Monthly c Occasionally c	
Diesel fumes	c	c	Daily c Weekly c	Monthly c Occasionally c	
Welding fumes	c	c	Daily c Weekly c	Monthly c Occasionally c	
Solvent fumes	c	c	Daily c Weekly c	Monthly c Occasionally c	
Oil / Gas well fumes	c	c	Daily c Weekly c	Monthly c Occasionally c	
Herbicides (to kill plants)	c	c	Daily c Weekly c	Monthly c Occasionally c	
Fungicides (to treat grain)	c	c	Daily c Weekly c	Monthly c Occasionally c	
Insecticides (to kill insects)	c	c	Daily c Weekly c	Monthly c Occasionally c	
Molds	c	c	Daily c Weekly c	Monthly c Occasionally c	
Radiation	c	c	Daily c Weekly c	Monthly c Occasionally c	
Other, Specify _____	c	c	Daily c Weekly c	Monthly c Occasionally c	

B-59 How often do you (did you) wear a dust mask when exposed to grain dust?

c Always c Most of the time c Sometimes c Never

B-60 We wish to find out more about respiratory health of rural people. Would you be willing to be contacted about having breathing and/or allergy tests at a nearby location?

c Yes

c No

c I would like more information

IF THERE IS ONLY ONE ADULT IN YOUR FAMILY, PLEASE SKIP TO THE LAST PAGE.

IF THERE IS ANOTHER ADULT IN YOUR FAMILY, PLEASE CONTINUE ON THE NEXT PAGE.

REMEMBER TO COMPLETE THE CONTACT INFORMATION ON THE LAST PAGE!
(THIS INFORMATION WILL BE REMOVED FROM YOUR QUESTIONNAIRE TO ENSURE CONFIDENTIALITY.)

SECTION B INDIVIDUAL QUESTIONS

ADULT 2

NOW, PLEASE ANSWER THE FOLLOWING QUESTIONS ABOUT ADULT # 2.

B-1 Age as of January 1st, 2010: _____

B-2 Date of birth: MM_____ DD____ YY_____

B-3 Sex: Male c Female c

B-4 Highest level of education:

- c Less than high school
- c Completed high school
- c Completed university
- c Completed post-secondary education other than above

B-5 What is your ethnic background?

- c Caucasian
- c First Nation
- c Metis
- c Other → **Please specify:** _____

B-6 What is your height? _____cm. **OR** _____ft and in.

B-7 What is your weight? _____Kg. **OR** _____lbs

B-8 What is your marital status? **(Please check only one)**

- c Married
- c Common law/living together
- c Widowed
- c Divorced/separated
- c Single, never married

RESPIRATORY HEALTH

COUGH

B-9 Do you usually have a cough?

- c Yes
- c No → **If no, go to question B-12.**

B-10 Do you usually cough like this on most days for 3 consecutive months or more during the year?

- c Yes
- c No

B-11 For how many years have you had this cough?
_____ years

PHLEGM

B-12 Do you usually bring up phlegm from your chest? c Yes

c No → **If no, go to question B-15.**

B-13 Do you bring up phlegm like this on most days for 3 consecutive months or more during the year?

- c Yes
- c No

B-14 For how many years have you had trouble with phlegm?
_____ years

WHEEZE

B-15 Does your chest ever sound wheezy or whistling:

	Yes	No
1. When you have a cold?	c	c
2. Apart from colds?	c	c
3. Most days or nights?	c	c

If YES to 1, 2, OR 3, for how many years has this been present? _____number of years

B-16 Have you ever had an attack of wheezing that has made you feel short of breath?

- c Yes
- c No

If YES, have you ever required medicine or treatment for the(se) attack(s)?

- c Yes
- c No

BREATHLESSNESS

B-17 Are you troubled by shortness of breath when hurrying on the level or walking up a slight hill?

- c Yes
- c No

B-18 Do you have to walk slower than people of your age because of breathlessness?

- c Yes
- c No

B-19 Do you ever have to stop for breath when walking at your own pace on the level?

- c Yes
- c No

B-20 Do you ever have to stop for breath after walking about 100 yards (or after a few minutes) on the level?

c Yes

c No

B-21 Are you too breathless to leave the house or breathless on dressing or undressing?

c Yes

c No

ASTHMA

B-22 Have you ever had asthma? c Yes

c No → **If no, go to question B-26.**

B-23 **If Yes to B-22:**

Do you still have it? c Yes c No

Was it confirmed by a doctor? c Yes c No

At what age did it start? ____ age in years

If you no longer have it, at what age did it stop? ____ age in years

B-24 **If yes to B-22**, how many times have you required services for asthma from the following places during the past 12 months?

Hospital inpatient: _____ times

Emergency room outpatient: _____ times

Doctor's office: _____ times

B-25 **If yes to B-22,** which of the following statements best describes your asthma medication use in the past 12 months:

c Never in the past 12 months
c At least once in the past 12 months
c At least once per month
c At least once per week
c Every day

ALLERGIES

B-26 Have you ever had an allergic reaction to any of the following: **(Please check all that apply).**

1. House dust	c Yes	c No
2. Cats	c Yes	c No
3. Dogs	c Yes	c No
4. Grasses	c Yes	c No
5. Pollens	c Yes	c No
6. Molds	c Yes	c No
7. Others,	c Yes	c No

Please specify:_____

PHYSICAL ACTIVITY

B-27 Do you exercise?

c Yes → **If yes, how many times a week?** _____times a week

c No → **If no, go to question B-29.**

B-28 How long do you usually exercise? c Less than 15 minutes

c 15 to 30 minutes c 31 to 60 minutes

c More than 60 minutes c Don't Know

B-29 In a **typical week** in the past **3 months**, how much time did you usually spend on a computer, including playing computer games and using the Internet or World Wide Web? **(Please do not include time spent at work or at school)**

c None
c Less than 1 hour
c From 1 to 2 hours
c From 3 to 5 hours
c From 6 to 10 hours
c From 11 to 14 hours
c From 15 to 20 hours
c More than 20 hours

B-30 In a **typical week** in the past **3 months**, how much time did you usually spend watching television or videos?

c None
c Less than 1 hour
c From 1 to 2 hours
c From 3 to 5 hours
c From 6 to 10 hours
c From 11 to 14 hours
c From 15 to 20 hours
c More than 20 hours

EARLY LIFE EXPOSURES

B-31 Have you ever lived on a farm?

c Yes

c No

c Don't know

B-32 Did you live on a farm during your first year of life?

c Yes → **If yes, what type of farm? (Check all that apply)**

c Grain
c Livestock

c No

c Don't know

B-33 Did your mother smoke while she was pregnant with you?

c Yes

c No

c Don't know

Adult 2

B-34 What was your birth weight?
_____ pounds or _____ grams

c Don't know

B-35 Were you breastfed as a child?

c Yes → **If yes, was it for 6 months or longer?** c Yes c No

c No

c Don't know

CIGARETTE SMOKING

B-36 Have you ever smoked cigarettes? **(If you have smoked less than 20 packs of cigarettes in your lifetime, answer no.)**

c Yes

c No → **If no, go to question B-43**

B-37 Do you now smoke cigarettes?

c Yes

c No

B-38 How old were you when you first started regular cigarette smoking? ____ years old

B-39 How many cigarettes do you smoke per day now?_____ cigarettes per day

B-40 On the average of the entire time you smoked, how many cigarettes did you smoke per day?
_____ cigarettes per day

B-41 If you have stopped smoking cigarettes completely, how old were you when you stopped?
_____ age stopped

B-42 If there have been periods when you abstained from smoking, indicate total years of abstinence from smoking. _____ years

B-43 Have you ever smoked a pipe regularly? **(Yes means more than 12 oz of tobacco in a lifetime)**

c Yes

c No

B-44 Have you ever smoked cigars regularly? **(Yes means more than 1 cigar a week for a year)**

c Yes

c No

B-45 Do you smoke a pipe or cigars regularly at present?

c Yes

c No

ALCOHOL CONSUMPTION

B-46 During the past 12 months, how often did you drink alcoholic beverages?

c Never
c Less than once a month
c Once a month
c 2 to 3 times a month
c Once a week
c 2 to 3 times a week
c 4 to 6 times a week
c Every day

B-47 How often in the past 12 months have you had 5 or more drinks on one occasion?

c Never
c Less than once a month
c Once a month
c 2 to 3 times a month
c Once a week
c More than once a week

MEDICAL HISTORY

B-48 In general would you say your health is:

c Excellent
c Very Good
c Good
c Fair
c Poor

B-49 During the past 12 months, were you seen by a doctor or other primary care giver for:

	Yes	No	Don't know
Stomach acidity or reflux?	c	c	c
An ear infection?	c	c	c
An injury?	c	c	c

B-50 Has a doctor or primary care giver ever said you have:

	Yes	No	Don't Know
Diabetes	c	c	c
Heart Disease	c	c	c
Heart Attack	c	c	c
Hardening of the arteries	c	c	c
High Blood Pressure	c	c	c
Cystic Fibrosis	c	c	c
Tuberculosis	c	c	c
Stroke	c	c	c
Cancer	c	c	c

If yes to cancer, please specify type(s):

_____

_____

Adult 2

CHEST ILLNESSES

B-51 Has a doctor ever said you had any of the following chest illnesses:

	Chest Illness	During the Past 12 Months		Ever In Your Life	
a.	Attack of bronchitis	c Yes	c No	c Yes	c No
b.	Pneumonia	c Yes	c No	c Yes	c No
c.	Hay Fever	c Yes	c No	c Yes	c No
d.	Sinus Trouble	c Yes	c No	c Yes	c No
e.	Chronic Bronchitis	c Yes	c No	c Yes	c No
f.	Emphysema	c Yes	c No	c Yes	c No
g.	COPD (Chronic Obstructive Pulmonary Disease)	c Yes	c No	c Yes	c No
h.	Sleep Apnea	c Yes	c No	c Yes	c No
i.	Other Chest Illness (Example chest operation) please specify:	c Yes	c No	c Yes	c No

B-52 **If yes to Chronic Obstructive Pulmonary Disease (COPD) in question B-51g,** how many times have you required services for COPD from the following places during the **past 12 months**?

Hospital inpatient: _____ times

Emergency room outpatient: _____ times

Doctor's office: _____ times

REST AND SLEEP

B-53 Do you snore?

c Yes

c No → **If no, go to question B-55.**

c Don't know

B-54 If you snore, is your snoring:

c Slightly louder than breathing?

c As loud as talking?

c Louder than talking?

c Very loud - can be heard in adjacent rooms?

B-55 How likely are you to doze off or fall asleep in the situations described below, in contrast to just feeling tired? This refers to your usual way of life in recent times. Even if you haven't done some of these things recently, try to work out how they would have affected you. **Please check one response choice for each situation.**

SITUATION	RESPONSE CHOICES			
	Would never doze	Slight chance of dozing	Moderate chance of dozing	High chance of dozing
Sitting and reading	c	c	c	c
Watching TV	c	c	c	c
Sitting inactive in a public place (e.g., a theatre or a meeting)	c	c	c	c
As a passenger in a car for an hour without a break	c	c	c	c
Lying down to rest in the afternoon when circumstances permit	c	c	c	c
Sitting and talking to someone	c	c	c	c
Sitting quietly after lunch without alcohol	c	c	c	c
In a car, while stopped for a few minutes in the traffic	c	c	c	c

FAMILY HISTORY

B-56 Have the following members of your biological family ever had:

	FATHER			MOTHER			BROTHER/SISTER		
	Yes	No	Don't Know	Yes	No	Don't Know	Yes	No	Don't Know
Diabetes	c	c	c	c	c	c	c	c	c
Heart Disease	c	c	c	c	c	c	c	c	c
Heart Attack	c	c	c	c	c	c	c	c	c
Hardening of the arteries	c	c	c	c	c	c	c	c	c
High Blood Pressure	c	c	c	c	c	c	c	c	c
Cystic Fibrosis	c	c	c	c	c	c	c	c	c
Tuberculosis	c	c	c	c	c	c	c	c	c
Stroke	c	c	c	c	c	c	c	c	c
Lung Trouble (Asthma,Emphysema, Chronic Bronchitis)	c	c	c	c	c	c	c	c	c
Cancer	c	c	c	c	c	c	c	c	c
If yes to cancer, please specify type(s):	_____			_____			_____		

OCCUPATIONAL HISTORY

B-57 Please list all full-time jobs at which you have worked for at least one year, starting with your present or most recent job. Please state the job title and business as specifically as possible. For example, 'mixed farming' instead of 'farming'.

Job Title	Business, Industry or Service	Total number of Years at job
e.g. Nurse	Health Care	10
e.g. Farmer	Mixed Farming	30

Adult 2

B-58 Have you ever been exposed to any of the following in the work place?

	No	Yes	If Yes, how often?		How many years?
Grain Dust	c	c	Daily c Weekly c	Monthly c Occasionally c	
Mine dust (e.g. potash, uranium) Specify _____	c	c	Daily c Weekly c	Monthly c Occasionally c	
Asbestos dust	c	c	Daily c Weekly c	Monthly c Occasionally c	
Wood dust	c	c	Daily c Weekly c	Monthly c Occasionally c	
Other dust Specify _____	c	c	Daily c Weekly c	Monthly c Occasionally c	
Livestock	c	c	Daily c Weekly c	Monthly c Occasionally c	
Smoke from stubble burning	c	c	Daily c Weekly c	Monthly c Occasionally c	
Diesel fumes	c	c	Daily c Weekly c	Monthly c Occasionally c	
Welding fumes	c	c	Daily c Weekly c	Monthly c Occasionally c	
Solvent fumes	c	c	Daily c Weekly c	Monthly c Occasionally c	
Oil / Gas well fumes	c	c	Daily c Weekly c	Monthly c Occasionally c	
Herbicides (to kill plants)	c	c	Daily c Weekly c	Monthly c Occasionally c	
Fungicides (to treat grain)	c	c	Daily c Weekly c	Monthly c Occasionally c	
Insecticides (to kill insects)	c	c	Daily c Weekly c	Monthly c Occasionally c	
Molds	c	c	Daily c Weekly c	Monthly c Occasionally c	
Radiation	c	c	Daily c Weekly c	Monthly c Occasionally c	
Other, Specify _____	c	c	Daily c Weekly c	Monthly c Occasionally c	

How often do you (did you) wear a dust mask when exposed to grain dust?

B-59

c Always c Most of the time c Sometimes c Never

B-60 We wish to find out more about respiratory health of rural people. Would you be willing to be contacted about having breathing and/or allergy tests at a nearby location?

c Yes

c No

c I would like more information

IF THERE ARE MORE THAN TWO ADULT FAMILY MEMBERS LIVING IN YOUR HOUSEHOLD,

PLEASE CONTINUE IN THE GREEN BOOKLET.

REMEMBER TO COMPLETE THE CONTACT INFORMATION ON THE LAST PAGE!

(THIS INFORMATION WILL BE REMOVED FROM YOUR QUESTIONNAIRE TO ENSURE CONFIDENTIALITY.)

CONTACT INFORMATION (PLEASE PRINT)

NAME:_____ Age: _____ c Male c Female
(Name of person completing the survey)

_____
Address (number and street and Box Number)

_____ , _____
Town Postal code

If you live on a farm give the land location of your residence.

_____
Land location (quarter, section, township, meridian)

Telephone Numbers **(check most preferred):**

Work _____ c

Home _____ c

Cell _____ c

THIS IS THE END OF THE SURVEY.

THANK YOU VERY MUCH FOR YOUR HELP!

Appendix- B: Example of Syntax from SPSS

```
* Generalized Estimating Equations.
GENLIN HRC_CANCERS_BASELINE (REFERENCE=FIRST) BY DIESEL_RECODED
(ORDER=DESCENDING)
/MODEL DIESEL_RECODED INTERCEPT=YES
DISTRIBUTION=BINOMIAL LINK=LOGIT
/CRITERIA METHOD=FISHER(1) SCALE=1 MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE)
SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95 LIKELIHOOD=FULL
/REPEATED SUBJECT=HOUSEID WITHINSUBJECT=PERSONID SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES
COVB=ROBUST MAXITERATIONS=100 PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
/MISSING CLASSMISSING=EXCLUDE
/PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION
(EXPONENTIATED).
```

Appendix- C: Certificate of Approval from Biomedical Research Ethics Board

UNIVERSITY OF SASKATCHEWAN

Biomedical Research Ethics Board (Bio-REB) 09/Mar/2020

Certificate of Approval

Application ID: 1789

Principal Investigator: Punam Pahwa

Department: Department of Community Health and Epidemiology

Locations Where Research Activities are Conducted: Department of Community Health & Epidemiology University of Saskatchewan, Canada

Student(s): Amal Khan

Funder(s):

Sponsor:

Title: Prevalence and Incidence of Hormonal-Related Cancers (Hrcs) in Rural Saskatchewan

Protocol Number:

Approved On: 06/Mar/2020

Expiry Date: 05/Mar/2021

Approval Of:
* Biomedical Application - Secondary Use of Health Data 01-Amal
* Variable List

Acknowledgment Of:
* Bio 1789 NER_response
* McMaster Certificate
* tcps2_core_certificate

Review Type: Delegated Review

IRB Registration Number: Not Applicable

CERTIFICATION

The University of Saskatchewan Biomedical Research Ethics Board (Bio-REB) has reviewed the above-named project. The project is acceptable on scientific and ethical grounds. The principal investigator has the responsibility for any other administrative or regulatory approvals that may pertain to this project, and for ensuring that the authorized project is carried out according to governing law. This approval is valid for the specified period provided there is no change to the approved project.

FIRST TIME REVIEW AND CONTINUING APPROVAL

The University of Saskatchewan Research Ethics Boards review above minimal risk projects at a full-board (face-to-face) meeting. If a project has been reviewed at a full board meeting, a subsequent project of the same protocol may be reviewed through the delegated review process. Any research classified as minimal risk is reviewed through the delegated (subcommittee) review process. The initial Certificate of Approval includes the approval period the REB has assigned to a study. The Status Report form must be submitted within one month prior to the assigned expiry date. The researcher shall indicate to the REB any specific requirements of the sponsoring organizations (e.g. requirement for full-board review and approval) for the continuing review process deemed necessary for that project.

REB ATTESTATION

In respect to clinical trials, the University of Saskatchewan Research Ethics Board complies with the membership requirements for Research Ethics Boards defined in Part 4 of the Natural Health Products Regulations and Part C Division 5 of the Food and Drug Regulations and carries out its functions in a manner consistent with Good Clinical Practices. Members of the Bio-REB who are named as investigators, do not participate in the discussion related to, nor vote on such studies when presented to the Bio-REB. This approval and the views of this REB have been documented in writing. The University of Saskatchewan Biomedical Research Ethics Board is constituted and operates in accordance with the current version of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2 2018).

Digitally Approved by Dr. Gordon McKay, Ph.D.
Chair, Biomedical Research Ethics Board
University of Saskatchewan

1 / 1