Socioeconomic Status and Risk of Chronic Respiratory Conditions in Rural Saskatchewan

A Thesis

Submitted to the College of Graduate and Post Doctoral 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, Canada

By

Francis T. Abayateye

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a graduate 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, 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.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Dean

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

Department Head

Department of Community Health and Epidemiology
Box 7, Health Science Building, 107 Wiggins Road
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5E5

Abstract

Chronic respiratory conditions (CRC) are a leading cause of morbidity and mortality globally. Research suggests that people of lower socioeconomic status (SES) are more likely to have CRC than those of higher status. However, the majority of these studies are cross-sectional in design, and in addition, have not considered the experience of rural dwellers, particularly in North America. Informed by a social determinants of health perspective, the primary purpose of this study was to: 1) determine the incidence of CRC in rural-dwelling adults in Saskatchewan; and 2) examine the association between SES and the incidence of CRC. The data source was the Saskatchewan Rural Health Study (SRHS), a prospective cohort study that consisted of a baseline survey in 2010 and a follow-up survey in 2014. The dependent variable was CRC, comprised of self-reported asthma and/or chronic obstructive pulmonary disease (COPD). The primary exposure of interest was SES, and assessed using measures of household income adequacy, educational attainment, and financial strain. Survival analysis was used to identify the risk factors of incidence of CRC, adjusting for various covariates. The cumulative incidence of CRC was 7.01%. Compared with high-income adequate participants, those with low and low-middle income adequacy had 2.22 times (95% CI: 1.01 – 4.89) and 1.66 times (95% CI: 1.08 – 2.56) higher risk of CRC, respectively. Financial strain and education were not related to the risk of CRC. Other statistically significant risk factors included the use of household natural gas, smoking status, parental history of lung disease, allergy, and diabetes. In conclusion, lower household income was associated with an increased risk of developing CRC over the four-year study period. I suggest that rural health policies should pay attention to the socioeconomic circumstances of rural people and not just access/distance to health services.

Acknowledgement

My deepest and sincere gratitude goes to my supervisor, Dr Bonnie Janzen, for critically examining my scripts and for his constructive suggestions, painstaking corrections, continuous encouragement, and unflinching support. I deemed it a privilege to learn under your supervision. I am also grateful to my advisory committee members, Dr Punam Pahwa and Dr Joshua Lawson, for their participation and for providing me with valuable suggestions throughout the writing of this thesis. A special thank you goes to my external Examiner, Dr Hug Mobinul for his insightful contribution to making this thesis better.

I would further like to express my gratitude to the University for providing me with financial assistance. I also wish to thank all the faculty, staff and graduate students in the Department of Community Health and Epidemiology for their generous support and encouragement necessary to complete this endeavour.

I want to thank my family and friends, especially my wife, Vera Amo Larbi, for their support and prayers.

Dedication

I dedicate this thesis to my little daughter Gwyneth N. K. Abayateye, and my wife, Vera Amo Larbi.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AB	Alberta
ACOS	Asthma COPD Overlap Syndrome
ATS	American Thoracic Society
BC	British Columbia
BMI	Body mass index
CCDSS	Canadian Chronic Disease Surveillance System
CDSS	Commission on Social Determinants of Health
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Diseases
CRC	Chronic Respiratory Conditions
CSDH	Commission on Social Determinants of Health
DAG	Directed Acyclic Graphs
DALY	disability-adjusted life year
EMR	Electronic medical registry
ERS	European Respiratory Society
FEV_1	Forced expiratory volume in 1 second
FVC	Forced vital capacity
GBD	Global burden of disease
GOLD	Global Initiative for Chronic Obstructive Pulmonary Disease
HR	Hazard ratio
MAR	Missing at random
MCAR	missing completely at random
MIZ	Metropolitan Influence Zone
MNAR	Missing not at random
NB	New Brunswick
NCD	Non-communicable disease
NL	Newfoundland and Labrador
NPHS	National Population Health Survey
OR	Odds ratio
Ref	Reference
SES	Socioeconomic Status
SK	Saskatchewan
SRHS	Saskatchewan Rural Health Study
US	United States
WHO	World Health Organization
YLD	Year-lived-with-disability

CHAPTER 1
INTRODUCTION

1.1. Background

Chronic respiratory conditions (CRC) are diseases of the airways and other structures of the lung [1] and are one of the leading causes of morbidity and mortality globally [2]. Among the most common CRC are chronic obstructive pulmonary disease (COPD), which includes emphysema and chronic bronchitis, and asthma[1] [3,4]. COPD is the third-leading cause of death worldwide [5] and the fifth-leading cause of death in Canada [6]. COPD is prevalent mainly among adults 35 years old and above and increases steadily with age [7]. Asthma, although most common in childhood [8], also occurs in adults, with global estimates of adult doctor-diagnosed asthma, clinical/treated asthma, and wheezing prevalence reported at 4.3%, 4.5%, and 8.6%, respectively [9]. In addition to resulting in profound levels of human loss and suffering, the economic burden of CRC is increasing globally, including in Canada [10,11].

CRC are the result of the interaction between genetic and environmental factors and, for this reason, are referred to as "complex diseases" [12]. Previous studies established several factors associated with increased risk of CRC. Risk factors most commonly associated with CRC include smoking, exposure to particulate matter, physical inactivity, diabetes, and cardiovascular disease [13–17]. Considerable research also suggests an association between CRC and socioeconomic status (SES) [18–22], with those of lower income and education, for example, more likely to report the presence of CRC when compared to their higher SES counterparts [18,23,24]. Population health perspectives characterize SES as a crucial determinant of health, shaping individuals' and communities' exposure to a complex range of health-enhancing or health-damaging life conditions throughout the life course [25].

[1] The range of diseases labeled as CRC vary in the literature, with some of the most common being chronic obstructive pulmonary disease (COPD), emphysema, chronic bronchitis, asthma, cystic fibrosis, sleep apnea, and occupational lung diseases [3,4]. However, in this study, CRC refers only to COPD and asthma, whereby COPD includes emphysema and chronic bronchitis.

1.2. Problem Statement

Although research evidence suggesting that low SES may play a role in the development of CRC is increasing, a number of limitations are evident. One of the most important limitations is that the majority of studies incorporating SES as exposure in relation to CRC have been cross-sectional [26–30]. Cross-sectional studies are among the weakest observational epidemiological designs because they are usually unable to detect whether the exposure of interest precedes or follows the disease; more specifically, whether low SES leads to CRC or vice versa. Thus, although numerous studies have linked SES with CRC, the use of prevalence results in a lack of clarity concerning whether SES is associated with disease development, duration, or some combination of the two.

Another limitation is that most studies of SES and CRC have not considered the experience of rural dwellers, particularly in North America. Compared to those in urban settings, rural residents have a demographic profile that may contribute to a higher incidence or prevalence of CRC, such as greater proportions of older people and those of Indigenous ancestry [31]. Also, some of the risk factors that may predispose individuals to CRC, such as smoking and particular occupational exposures, may be more prevalent among rural dwellers than in the general population [32,33]. More importantly, due to urban/rural differences in age structure, occupation, and/or educational opportunities, typical measures of SES may be less indicative of *power, prestige and access to resources* in rural dwellers — fundamental characteristics believed to underlie the development of health inequities [34–37]. Finally, with a few exceptions, the vast majority of studies of SES and CRC in rural settings are cross-sectional in design [27,38], leading to the interpretive challenges described above.

1.3. Study Aim and Research Questions

The overall objective of this study was to examine the longitudinal association between SES and CRC among rural-dwelling adults in Saskatchewan. This study addressed two research questions:

1. What is the incidence of CRC among rural-dwelling Canadians?
2. Is SES associated with incident CRC in rural Canadians?

1.4. Significance of Study

Research examining the association between SES and CRC among rural dwellers has several potential benefits, yet such studies, particularly longitudinal ones, have rarely been conducted in a North American setting. A longitudinal study allows for the estimation of incidence and is able to better demonstrate the temporal relationship between exposure and outcome. Rural populations have some unique characteristics compared to urban populations [39]. Therefore, it is important to examine whether relationships between SES and CRC, usually observed in general population samples, also hold true in a rural context. In addition to contributing to the scholarly literature, the results of this study may identify subgroups of rural people whose risk of CRC could be heightened by their socioeconomic circumstances and, thus, can inform the development of tailored interventions.

CHAPTER 2
LITERATURE REVIEW

2.1. Introduction

This chapter begins with the definition, diagnosis, measurement and descriptive epidemiology of CRC. Risk factors for CRC are described, highlighting SES, and the conceptual framework which informs the study is presented. The chapter ends with a critical appraisal of the literature on SES and CRC, providing the rationale for the present study.

2.2. Definition, Diagnosis, and Measurement of Chronic Respiratory Conditions

The label CRC may be used in reference to any number of diseases that affect the lungs and airways [40]. However, this study will only be concerned with COPD and asthma, as they are the most common obstructive lung diseases worldwide [41]. Further, even though chronic bronchitis and emphysema may be clinically different from COPD, they are generally considered components of COPD[7] and will be treated as such in this study.

There exist several definitions for COPD, without any single definition having a clear superiority over the other. The varying and improved definitions over time reflect not only the complexity of the disease but also increasing knowledge. Earlier definitions came from the work of the American Thoracic Society (ATS) [42] and the European Respiratory Society (ERS) [43]. The Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD) also came up with their own definition in 2001, with revisions in 2006 and 2011. According to their 2011 recommendations, COPD, [44]

> *"a common preventable and treatable disease is characterized by persistent airflow limitation that is usually progressive and associated with an enhanced chronic inflammatory response in the airways and the lung to noxious particles or gases. Exacerbations and comorbidities contribute to the overall severity in individual patients."*

A clinical diagnosis of COPD, according to GOLD, requires information on patients' symptoms, knowledge of previous exposures, and the use of spirometry[44]:

> *A clinical diagnosis of COPD should be considered in any patient who has dyspnea, chronic cough and/or sputum production, and a history of exposure to risk factors for the disease. Spirometry is required to make the*

diagnosis in this clinical context; the presence of a post-bronchodilator FEV_1/FVC less than 0.70 confirms the presence of persistent airflow limitation and thus, of COPD.

Another commonly recommended cut-off to indicate airway obstruction indicative of COPD is a lower limit of normal FEV_1/FVC [45]. COPD is often misdiagnosed and underdiagnosed in primary care due, in part, to inconsistent use of spirometry and application of clinical guidelines, [46–49], with some studies indicating as much as 72% to 93% of COPD cases being underdiagnosed [46]. In Canada, Hill et al. reported 32% undiagnosed and 30% misdiagnosed cases among individuals with COPD [47].

To obtain population estimates of COPD, studies depend mainly on doctor-diagnosed self-reported COPD and medical records, and sometimes, spirometry [18,27,28,50,51]. Different study methodologies and diagnostic criteria result in varied estimates of the prevalence and incidence of COPD [52–54]. For example, a Canadian study reported COPD prevalence using measured airflow obstruction to range between 5% to 15%, and these estimates are two to six times higher than self-reported measures of COPD prevalence [52]. Epidemiological studies rely mainly on field measurements of lung function using spirometry and simple questions to exclude asthma and other respiratory diseases [44].

The WHO defines asthma as a long-term condition which causes narrowing of the passage in the lungs due to inflammation and tightening of the muscles around the small airways, leading to symptoms of cough, wheeze, shortness of breath and chest tightness [55]. Given the emerging unanimity in the research community that asthma is an umbrella term for several diseases with similar clinical manifestations but different underlying pathophysiological mechanisms [56,57], a unified definition may be inappropriate.

Asthma diagnosis involves the identification of episodic symptoms and at least partially reversible airflow obstruction [58]. These “key points” are usually recommended by national health institution guidelines as there is currently no gold standard for asthma diagnosis. A spirometry measurement before and after administration of a short-acting bronchodilator is required for evidence of reversibility [58]. Subjective symptoms such as clinical symptoms, patient’s response to medication, and clinical history can also help diagnose asthma since objective criteria are not usually carried out in an office setting [59].

Various definitions are employed in epidemiological studies of asthma. As is the case with COPD, the different ways asthma is defined and measured may substantially impact the estimation of prevalence, incidence, and association with risk factors [60,61]. In a review of 122 published epidemiological studies of childhood asthma, 60 different operational definitions of asthma were identified [62]. Van Wonderen et al. observed that agreement between four seemingly very similar and commonly used definitions of asthma was relatively low and that, depending on the definition used, about one-third of the participants could be classified as "control" instead of "asthma cases" [62]. Generally, there are three main definitions of asthma used in epidemiological research: (1) self-reported doctor-diagnosed asthma, which is usually based on the question "*Have you ever been diagnosed with asthma*?" ; (2) clinical asthma, based on doctor-diagnosed asthma and/or an affirmative response to either of "*Have you ever been treated for asthma*?" or "*Have you been taking any medications or treatment for asthma during the last 2 weeks?*"; and (3) symptoms of asthma, based on doctor-diagnosed asthma, clinical asthma and/or an affirmative answer to "*During the last 12 months have you experienced attacks of wheezing or whistling breath?*" [9,63–65]. Researchers also sometimes depend on population-based hospital registries and data on anti-asthmatic medication from prescription registries [66].

Asthma and chronic obstructive pulmonary disease (COPD) are not always easy to differentiate, with older adult current smokers likely to have both COPD and asthma – a condition referred to as asthma-COPD overlap syndrome (ACOS) [67,68]. ACOS, like its components, do not have a unique definition [69,70]. The diagnostic guideline depends on the population from which the patients emerged [71]. For instance, in a longitudinal study involving patients suspected to have COPD and asthma visiting a hospital for the first time, ACOS was defined as "being diagnosed with COPD, positive bronchodilator test with >12% and 200 mL gain in FEV_1, and presence of clinical characteristics of asthma (previous history of asthma/wheezing)" [72]. In a population-based study using an electronic medical registry (EMR) and questionnaire, ACOS was defined as "the presence of COPD and asthma in the EMR, questionnaires and lung function or a combination of two or all of them" [73].

2.3. Burden and Economic Impact of CRC

Poor health affects the well-being of the individual involved and the resources and economic growth of the society and nation in which the person lives. Not only will human

resources not be available to work, but the government would also need to provide resources and income support for such people. Thus, when more people are sick in a nation, less work is done, and expenditure increases.

2.3.1. International

According to the 2017 Global Burden of Disease (GBD) study, non-communicable diseases (i.e. diseases that are not directly transmissible from one person to another) account for about 73% of mortality globally, with COPD and asthma contributing significantly to the rising burden of non-communicable diseases worldwide [2,74]. The Global Alliance against Chronic Respiratory Diseases was established by WHO with the sole aim of reducing the burden of chronic respiratory diseases, focusing particularly on the needs of people with CRC in low-income and middle-income countries [75]. The GBD study estimated, in 2017, a CRC all-age prevalence of about 544.9 million cases, an incidence of about 62.2 million new cases, and a year-lived-with-disability (YLD) of approximately 44.3 million [76]. From 2007 to 2017, the total number of deaths from CRC increased globally by 15.8% [74]. The disability-adjusted life year (DALY) (i.e. the sum of potential life lost due to premature mortality and the years of productive life lost due to disability) of CRC in 2017 was 112 million, an increase of 14.5% from the 2007 estimate [77]. The DALY change is higher in females (18.9%) than males (10.7%) [77]. High-income regions in the GBD study experienced the highest CRC prevalence at 10.6% in 2017 [4]; in contrast, Sub-Sahara Africa and South Asia experienced the lowest prevalence of 5.1% and 5.5%, respectively 2017 [4]. COPD and asthma are the primary CRC driving the surging burden of non-communicable diseases (NCD) [5].

COPD is the third leading cause of death globally and is responsible for about 6% of total deaths [5]. COPD is the most common cause of chronic respiratory disease-attributable deaths, accounting for approximately 41.9 deaths per 100,000 people in 2017. [4]. According to WHO, more than 200 million people suffer from COPD, with 65 million having moderate or severe airway disease. In 2017, the GBD study estimated a DALY of 81.6 million, with females experiencing the greatest consequences [77]. The change in DALY from 2007 to 2017 was 21.2% in females and 13.5% in males [77]. A systematic review and meta-analysis of 194 articles estimated COPD prevalence of about 9% in males and 6% in females, with significant differences across subregions [78].

Globally, asthma prevalence varies across regions and is on the rise mainly due to the increase in previously less prevalent regions such as Africa, Latin America, and some parts of Asia [79]. The causes of this rise remain elusive and a subject of speculation [79]. Prevalence also varies by sex, with adult females having a higher prevalence than adult males (9.6% verses 6.3%) [80]. The GBD study estimates of asthma prevalence varied between published reports: 220.4 million in 2000 [81], 327.1 million in 2005 [82], 334.2 million in 2010 [83], 241.7 million in 2013 [84], 358.2 million in 2015 [85], 339.4 million in 2016 [82], and 272.7 million in 2017 [76]. Death from asthma also increased significantly from 218,000 in the year 2000 to 495,100 in 2017, [74,81]. The GBD study also reported a disability-adjusted life year (DALY) of 22.8 million for asthma, with an increase of 9.7% between 2007 and 2017 only among females [77].

The economic burden of CRC is increasing globally. Generally, the economic burden of disease can be classified into direct and indirect costs [86]. Costs associated with utilizing health care resources for diagnosis and treatment of disease can be considered direct costs, while costs associated with reduced work productivity due to illness can be considered as indirect [87]. In 2019, an annual cost of about €380 billion was attributed to the care of patients with chronic respiratory diseases among 28 EU member states [10]. This cost includes direct primary and hospital healthcare, which amounts to at least €55 billion, lost production of at least €42 billion, and the monetized value of disability-adjusted-life-years (DALYs) of at least €280 billion [10]. A recent review of studies indicated that the direct cost of asthma varied from $150 per person-year (py) in the UAE to $3000/py in the US and indirect cost of about $1274/py in the Republic of Korea and that of COPD varied from $536/py to $4528/py [87]. The wide variation in these costs can be attributed to the difference in methodologies used in the cost-of-illness studies, making it difficult to associate differences in reported costs with the difference in the actual burden of COPD and asthma [88]. Nevertheless, these data highlight the continuous economic burden that CRC imposes across the globe.

2.3.2. Canada

Researchers have used several data sources to provide estimates of the prevalence and incidence of CRC in Canada, each with their own strengths and limitations [48]. The most common include government surveys of the general population, such as the Canadian Community Health Survey [89] and the Canadian Health Measures Survey [52], the latter of which includes both self-

reported doctor-diagnosed CRC and measures of lung function for a subgroup of the sample. Various research teams across Canada collect their own data using various methods [54]. Health care utilization records provide another source; for example, the Canadian Chronic Disease Surveillance System (CCDSS) produces population-based chronic disease prevalence and incidence estimates for COPD and asthma using administrative health data. The CCDSS covers 97% of the Canadian population and includes individuals who had any form of interaction with the health care system; COPD and asthma are identified based on physician diagnosis in an office or hospital setting [7]. The estimates provided below according to person, time and place characteristics are based on CCDSS data.

Figure 2.1 shows the prevalence of diagnosed asthma in Canada by age and sex. Asthma prevalence increases with age during childhood, with the highest prevalence reported for 10-14 years old males (22.2%) and 15-19 years old females (17.0%). Among adults, asthma prevalence is greatest among 20-24 years old, both males (15.8%) and females (14.2%), and then declines up until 30-34 years of age, remaining relatively stable until 60-64 years of age, after which prevalence increases slightly. One significant observation in Figure 2.1 is that the prevalence of asthma is higher for males than females during childhood and young adulthood, but then a cross-over is observed starting at 25-29 years of age, when the prevalence for males became lower than for females, continuing through to the older age group [7]. The incidence of asthma shows a similar age-sex pattern as prevalence, but the incidence rates peak between 1 and 4 years of age and the sex cross-over in incidence rates occurs between 15 and 19 years (Figure 2.4).

Figure 2.7 shows the prevalence of diagnosed COPD for Canadians aged 35 years and older, based on CCDSS data. The proportion of Canadians diagnosed with COPD increases steadily with age, with the highest prevalence among those 85 years and older. Sex/gender differences in COPD prevalence are minimal, up until the 60-64 age group, after which it becomes and remains consistently higher among males. As illustrated in Figure 2.12, the COPD incidence rate in Canada also increases steadily with age; men have a consistently higher COPD incidence than females in all age groups and this gap increases with age. However, incidence rate in general decreased over time for both male and female (Figure 2.11).

Data indicate that between 2000/01 and 2011/12, the age-standardized prevalence of both COPD and asthma increased in Canada[7]. For asthma (Figure 2.3), prevalence rose from 6.5% in 2000/01 to 10.8% in 2011/12. Similarly, the prevalence of COPD increased from 7% to 9.4%

among all Canadians aged 35 years and older (Figure 2.9). Females experienced a much higher relative increase in prevalence (42.8%) than males (22.9%) [7]. In contrast to prevalence, the age-adjusted incidence of COPD and asthma decreased during the same time period. Asthma age-standardized incidence rate declined from 8.94 per 1000 population in 2000 to 4.74 per 1000 population in 2012, a relative decline of about 46% (Figure 2.5). Similarly, the overall age-standardized incidence rate of COPD declined from 12.03 per 1000 population to 8.81 per 1000 population, a relative decline of approximately 26% [53]; females experienced a relative decline of about 22%, and males, 28% [7].

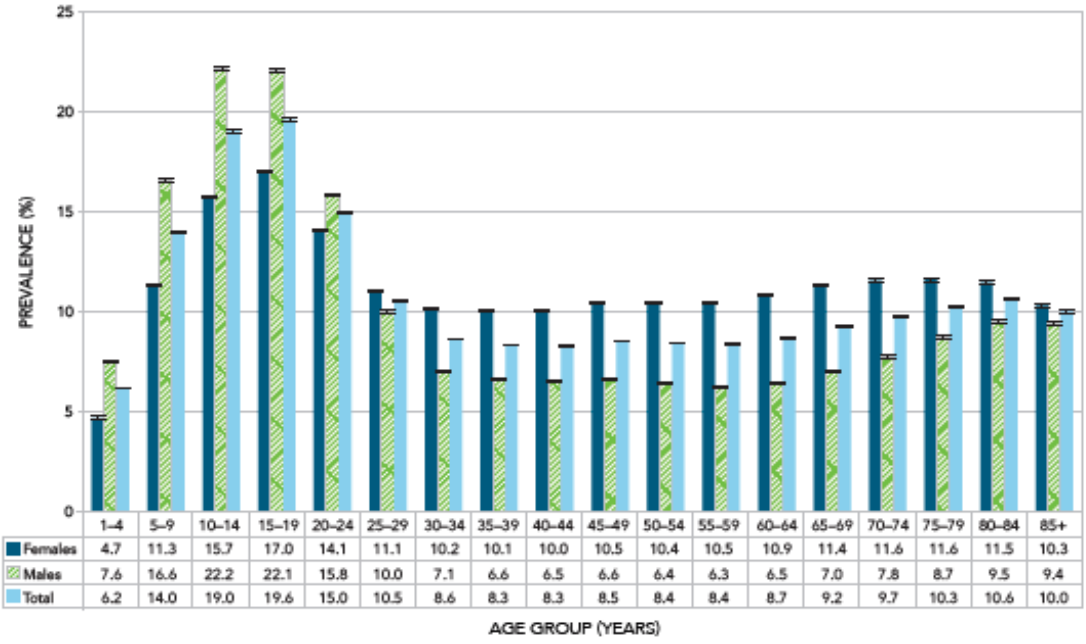
The age-standardized prevalence/incidence of COPD and asthma also varies among provinces and territories in Canada. The highest prevalence of asthma is observed in Ontario and Nova Scotia, with both exceeding the national estimate as a whole (Figure 2.2). The lowest proportion of people with asthma are observed in the territories. The highest incidence of asthma is reported in Yukon, Manitoba, and Ontario, which exceeds the national estimate, and the lowest in New Brunswick, the Northwest Territories, and Nunavut (Figure 2.6). Regarding COPD, the highest prevalence is in the Northwest Territories and Nova Scotia (Figure 2.8). The highest COPD incidence rates were recorded in the Northwest Territories and the lowest in Ontario, Newfoundland and Labrador, and Quebec (Figure 2.10)

Differences in COPD and asthma prevalence/incidence also exist between urban and rural populations [90]. Estimates of asthma prevalence among adults aged 20-44 in Canadian cities range from 4.4% to 6.3% for males and 5.2% to 9.5% for females [91]. Asthma prevalence in rural Canada (SK) is 8.6%, with a higher prevalence among females [92]. In a national longitudinal sample of adolescent Canadians, Lawson et al. estimated the incidence of asthma at 10.7 per 1000 person-years among urban dwellers and 6.4 per 1000 person-years among rural dwellers. Similarly, COPD prevalence in urban Canada ranges from 5.3% to 19.1%, depending on the urban center and operational definition [93]. COPD prevalence among rural dwellers ranges from 7.3% to 8.6% [39].

Canada's economic burden of chronic respiratory conditions comes from hospitals, physicians, drug costs, private institutions, and home care. All these costs are considered direct costs. Patra et al. estimated a total direct cost of $3.87 billion to the Canadian health care system due to chronic respiratory disease in 1998 [11]. Indirect costs due to CRC, such as productivity loss

due to premature death and disability, amounted to $5.67 billion, resulting in a total economic cost of $9.53 billion [11]. This cost is equivalent to $295 per capita [11].

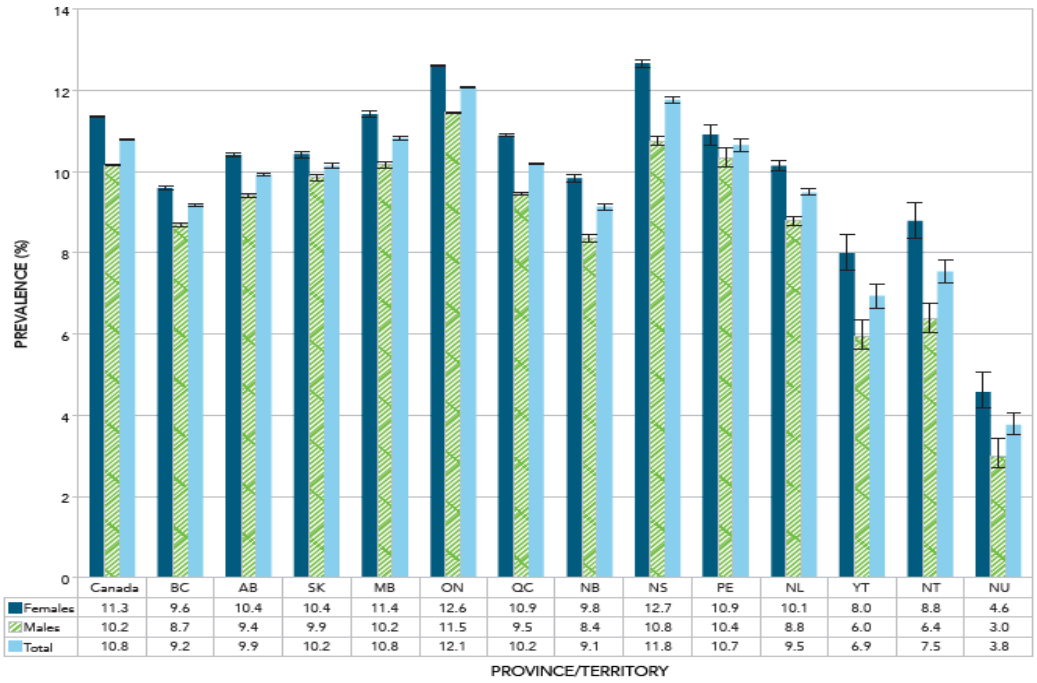
The Public Health Agency of Canada estimated direct and indirect costs due to respiratory system diseases [94]. The agency estimated a direct cost of $6.5 billion, which is equivalent to 6% of the total direct cost of all ills in Canada, and an indirect cost (lost productivity due to morbidity) of $3.1 billion, which is equivalent to 22% of the total indirect cost of all ills in Canada [94]. In the GOLD stage 2 and 3 study, 285 patients with an established diagnosis of moderate to severe COPD with a mean age of 70.4 were recruited from 23 sites across Canada [50]. The average annual COPD-related cost for these participants was $4,147 (2009 CAD), with 71% of the cost attributed to medication [50].



	1–4	5–9	10–14	15–19	20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69	70–74	75–79	80–84	85+
Females	4.7	11.3	15.7	17.0	14.1	11.1	10.2	10.1	10.0	10.5	10.4	10.5	10.9	11.4	11.6	11.6	11.5	10.3
Males	7.6	16.6	22.2	22.1	15.8	10.0	7.1	6.6	6.5	6.6	6.4	6.3	6.5	7.0	7.8	8.7	9.5	9.4
Total	6.2	14.0	19.0	19.6	15.0	10.5	8.6	8.3	8.3	8.5	8.4	8.4	8.7	9.2	9.7	10.3	10.6	10.0

Figure 2. 1. Prevalence of diagnosed asthma among Canadians aged one year and older, by age group and sex, Canada, 2011-2012

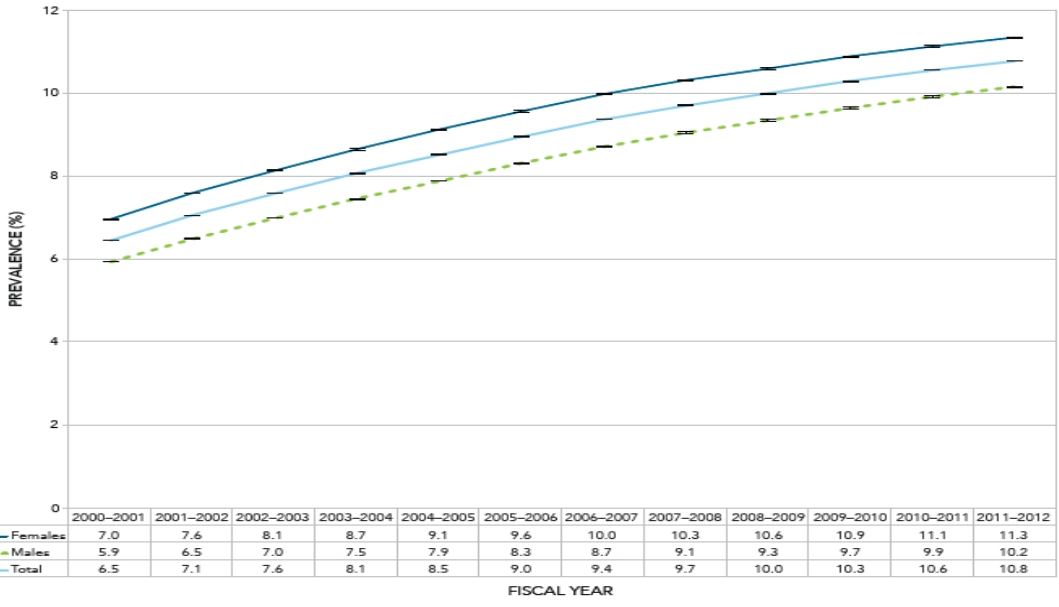
Source: Public Health Agency of Canada (2018)[7]



	Canada	BC	AB	SK	MB	ON	QC	NB	NS	PE	NL	YT	NT	NU
Females	11.3	9.6	10.4	10.4	11.4	12.6	10.9	9.8	12.7	10.9	10.1	8.0	8.8	4.6
Males	10.2	8.7	9.4	9.9	10.2	11.5	9.5	8.4	10.8	10.4	8.8	6.0	6.4	3.0
Total	10.8	9.2	9.9	10.2	10.8	12.1	10.2	9.1	11.8	10.7	9.5	6.9	7.5	3.8

Figure 2. 2. Age-standardized prevalence of diagnosed asthma among Canadians aged one year and older, by sex and province/territory, Canada, 2011-2012

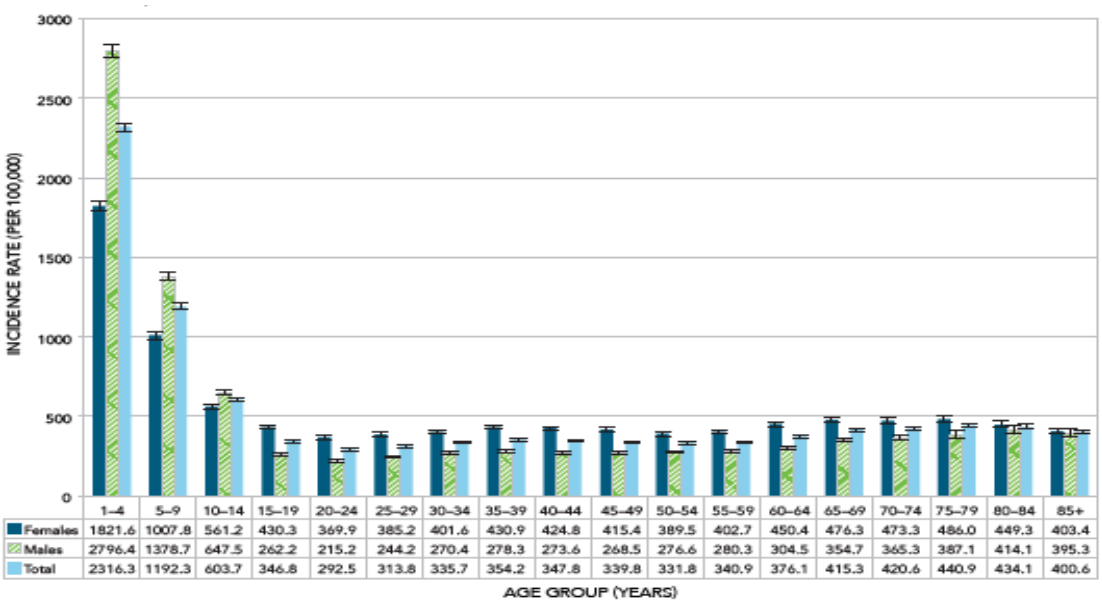
Source: Public Health Agency of Canada (2018)[7]



	2000–2001	2001–2002	2002–2003	2003–2004	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009	2009–2010	2010–2011	2011–2012
Females	7.0	7.6	8.1	8.7	9.1	9.6	10.0	10.3	10.6	10.9	11.1	11.3
Males	5.9	6.5	7.0	7.5	7.9	8.3	8.7	9.1	9.3	9.7	9.9	10.2
Total	6.5	7.1	7.6	8.1	8.5	9.0	9.4	9.7	10.0	10.3	10.6	10.8

Figure 2. 3. Age-standardized prevalence of diagnosed active asthma among Canadians aged one year and older, by sex and year, Canada, 2000–2001 to 2011–2012

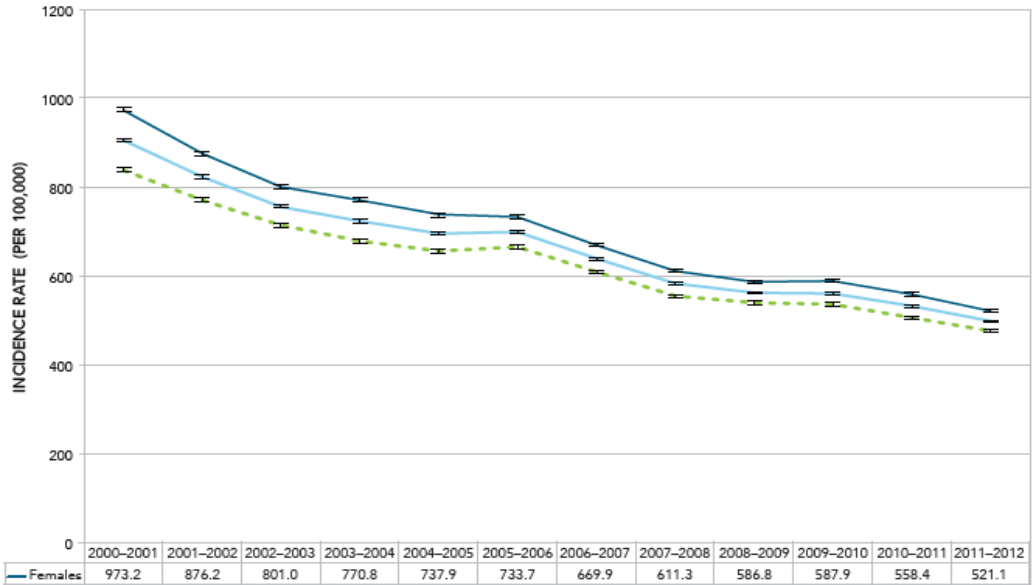
Source: Public Health Agency of Canada (2018)[7]



	1–4	5–9	10–14	15–19	20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69	70–74	75–79	80–84	85+
Females	1821.6	1007.8	561.2	430.3	369.9	385.2	401.6	430.9	424.8	415.4	389.5	402.7	450.4	476.3	473.3	486.0	449.3	403.4
Males	2796.4	1378.7	647.5	262.2	215.2	244.2	270.4	278.3	273.6	268.5	276.6	280.3	304.5	354.7	365.3	387.1	414.1	395.3
Total	2316.3	1192.3	603.7	346.8	292.5	313.8	335.7	354.2	347.8	339.8	331.8	340.9	376.1	415.3	420.6	440.9	434.1	400.6

Figure 2. 4. Figure 1. 1. Incidence rates of diagnosed asthma among Canadians aged one year and older, by age group and sex, Canada, 2011–2012

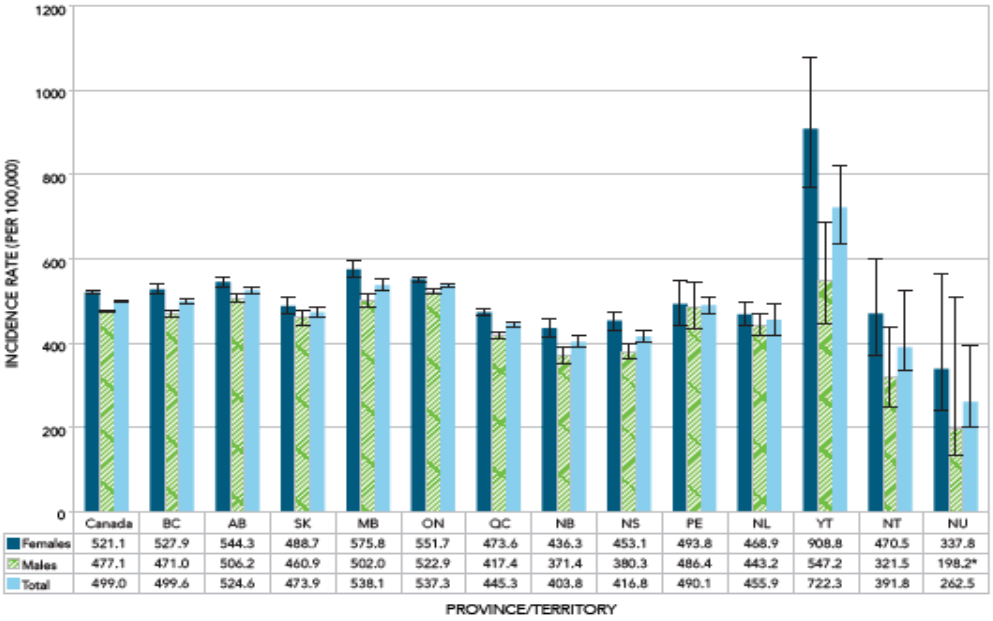
Source: Public Health Agency of Canada (2018)[7]



	2000–2001	2001–2002	2002–2003	2003–2004	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009	2009–2010	2010–2011	2011–2012
Females	973.2	876.2	801.0	770.8	737.9	733.7	669.9	611.3	586.8	587.9	558.4	521.1
Males	838.5	771.2	713.6	678.6	655.5	666.0	608.7	555.1	539.1	535.6	506.5	477.1
Total	904.5	822.4	756.1	723.4	695.7	698.8	638.5	582.6	562.5	561.4	532.2	499.0

Figure 2. 5. Age-standardized incidence of diagnosed asthma among Canadians aged one and above by sex and year, Canada, 2000-2001 to 2011-2-12

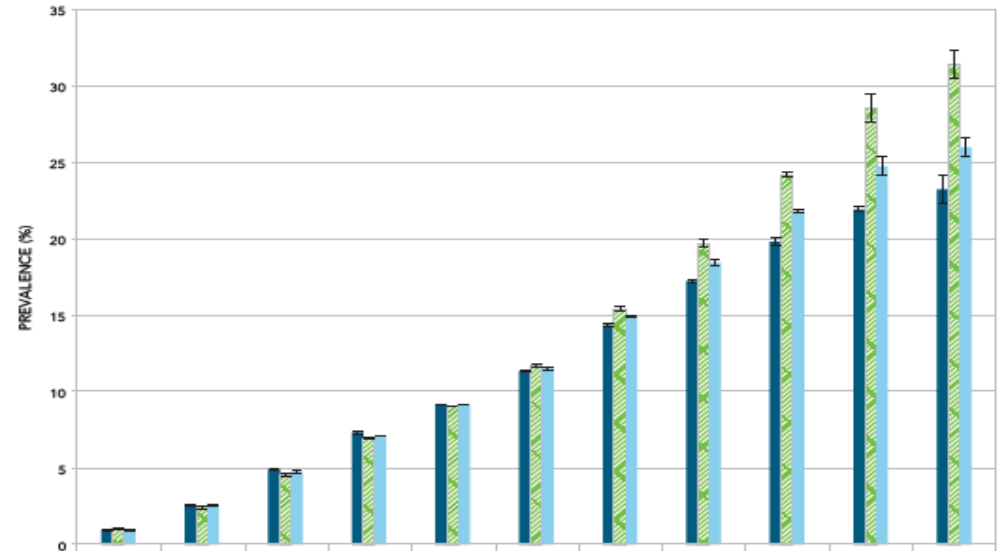
Source: Public Health Agency of Canada (2018)[7]



	Canada	BC	AB	SK	MB	ON	QC	NB	NS	PE	NL	YT	NT	NU
Females	521.1	527.9	544.3	488.7	575.8	551.7	473.6	436.3	453.1	493.8	468.9	908.8	470.5	337.8
Males	477.1	471.0	506.2	460.9	502.0	522.9	417.4	371.4	380.3	486.4	443.2	547.2	321.5	198.2*
Total	499.0	499.6	524.6	473.9	538.1	537.3	445.3	403.8	416.8	490.1	455.9	722.3	391.8	262.5

Figure 2. 6. Age-standardized incidence rates of diagnosed asthma among Canadians aged one year and older, by sex and province/territory, Canada, 2011-2012

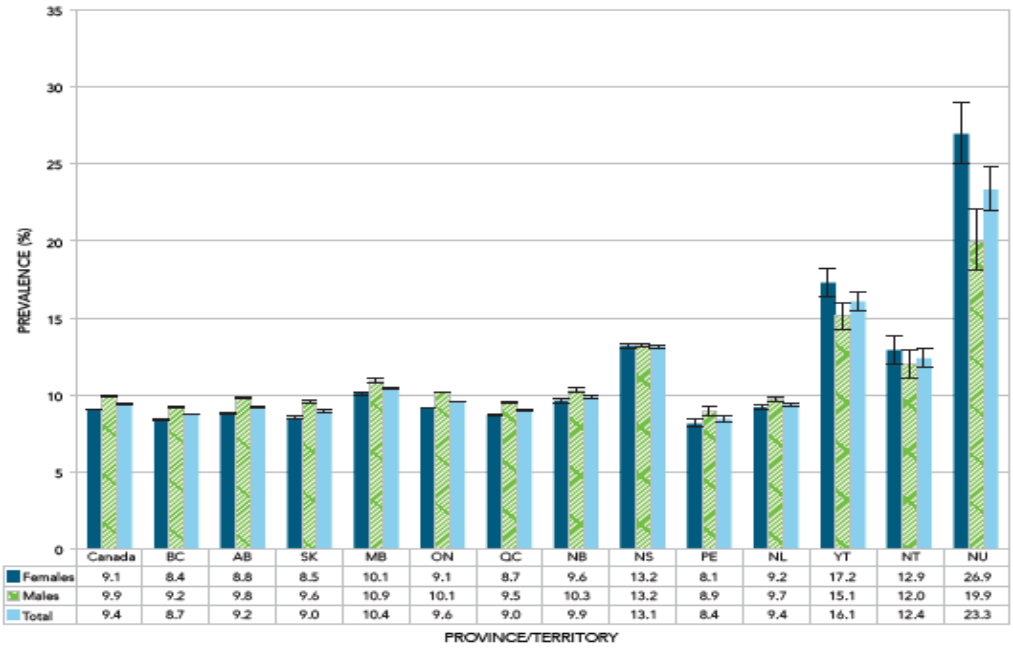
Source: Public Health Agency of Canada (2018)[7]



	35–39	40–44	45–49	50–54	55–59	60–64	65–69	70–74	75–79	80–84	85+
Females	0.9	2.5	4.9	7.3	9.1	11.3	14.3	17.2	19.8	21.9	23.2
Males	1.0	2.4	4.5	6.9	9.0	11.7	15.4	19.7	24.2	28.5	31.4
Total	0.9	2.5	4.7	7.1	9.1	11.5	14.9	18.4	21.8	24.7	26.0

Figure 2. 7. Prevalence of diagnosed COPD among Canadians aged 35 years and older, by age group and sex, Canada, 2011–2012

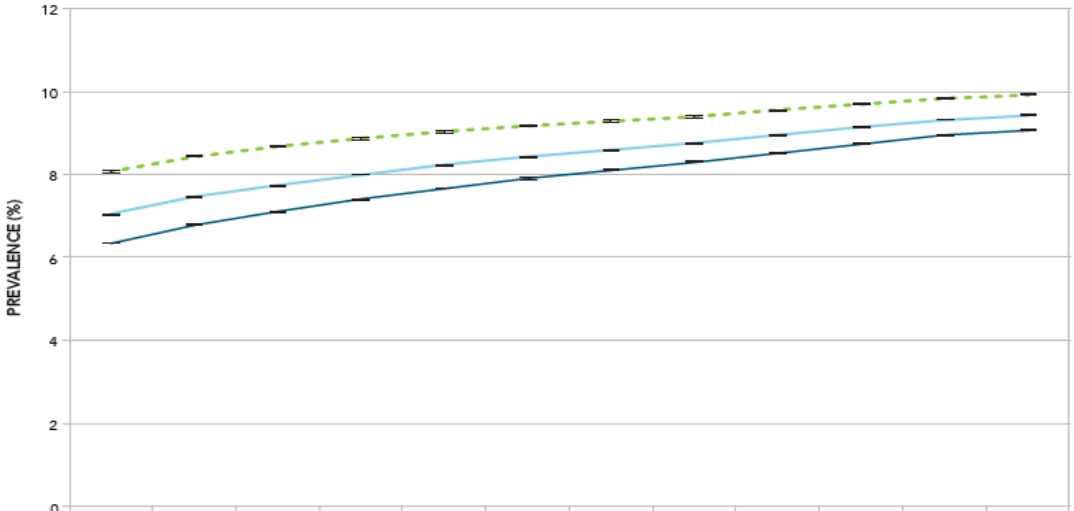
Source: Public Health Agency of Canada (2018)[7]



	Canada	BC	AB	SK	MB	ON	QC	NB	NS	PE	NL	YT	NT	NU
Females	9.1	8.4	8.8	8.5	10.1	9.1	8.7	9.6	13.2	8.1	9.2	17.2	12.9	26.9
Males	9.9	9.2	9.8	9.6	10.9	10.1	9.5	10.3	13.2	8.9	9.7	15.1	12.0	19.9
Total	9.4	8.7	9.2	9.0	10.4	9.6	9.0	9.9	13.1	8.4	9.4	16.1	12.4	23.3

Figure 2. 8. Age-standardized prevalence of diagnosed COPD among Canadians aged 35 years and older, by sex and province/territory, Canada, 2011–2012

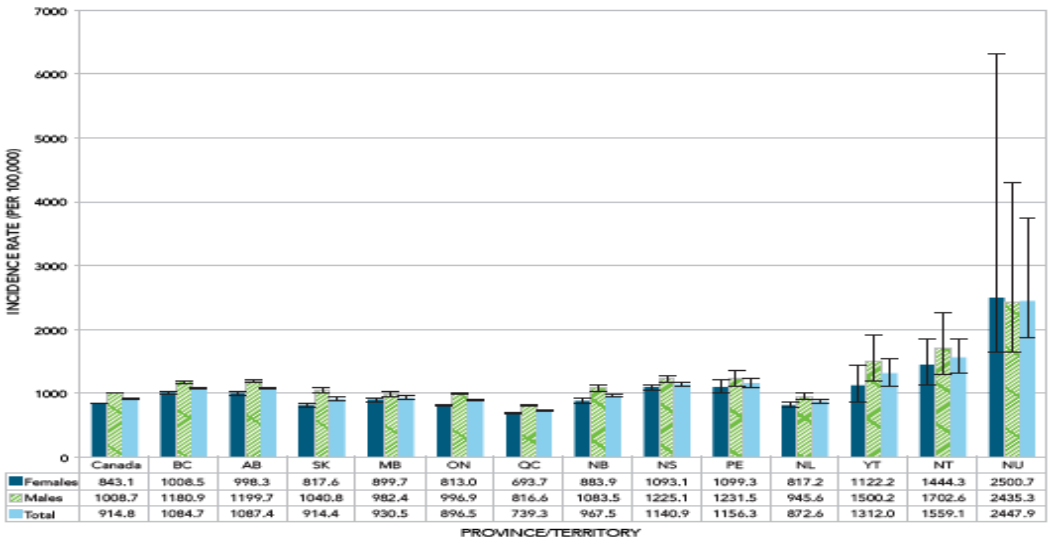
Source: Public Health Agency of Canada (2018)[7]



	2000–2001	2001–2002	2002–2003	2003–2004	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009	2009–2010	2010–2011	2011–2012
Females	6.3	6.8	7.1	7.4	7.7	7.9	8.1	8.3	8.5	8.7	8.9	9.1
Males	8.1	8.4	8.7	8.9	9.0	9.2	9.3	9.4	9.5	9.7	9.8	9.9
Total	7.0	7.5	7.7	8.0	8.2	8.4	8.6	8.7	8.9	9.1	9.3	9.4

Figure 2. 9. Age-standardized prevalence of diagnosed COPD among Canadians aged 35 years and older, by sex and year, Canada, 2000–2001 to 2011–2012

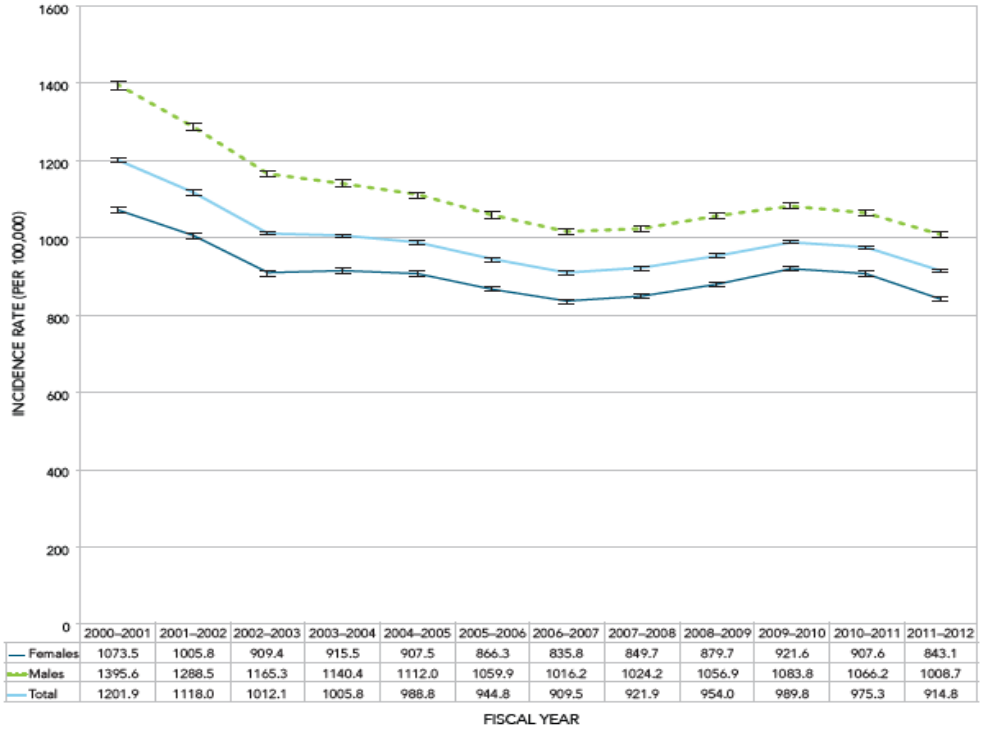
Source: Public Health Agency of Canada (2018)[7]



	Canada	BC	AB	SK	MB	ON	QC	NB	NS	PE	NL	YT	NT	NU
Females	843.1	1008.5	998.3	817.6	899.7	813.0	693.7	883.9	1093.1	1099.3	817.2	1122.2	1444.3	2500.7
Males	1008.7	1180.9	1199.7	1040.8	982.4	996.9	816.6	1083.5	1225.1	1231.5	945.6	1500.2	1702.6	2435.3
Total	914.8	1084.7	1087.4	914.4	930.5	896.5	739.3	967.5	1140.9	1156.3	872.6	1312.0	1559.1	2447.9

Figure 2. 10. Age-standardized incidence rates of diagnosed COPD among Canadians aged 35 years and older, by sex and province/territory, Canada, 2011–2012

Source: Public Health Agency of Canada (2018)[7]



	2000–2001	2001–2002	2002–2003	2003–2004	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009	2009–2010	2010–2011	2011–2012
Females	1073.5	1005.8	909.4	915.5	907.5	866.3	835.8	849.7	879.7	921.6	907.6	843.1
Males	1395.6	1288.5	1165.3	1140.4	1112.0	1059.9	1016.2	1024.2	1056.9	1083.8	1066.2	1008.7
Total	1201.9	1118.0	1012.1	1005.8	988.8	944.8	909.5	921.9	954.0	989.8	975.3	914.8

Figure 2. 11. Age-standardized incidence rates of diagnosed COPD among Canadians aged 35 years and older, by sex and year, Canada, 2000–2001 to 2011–2012

Source: Public Health Agency of Canada (2018)[7]

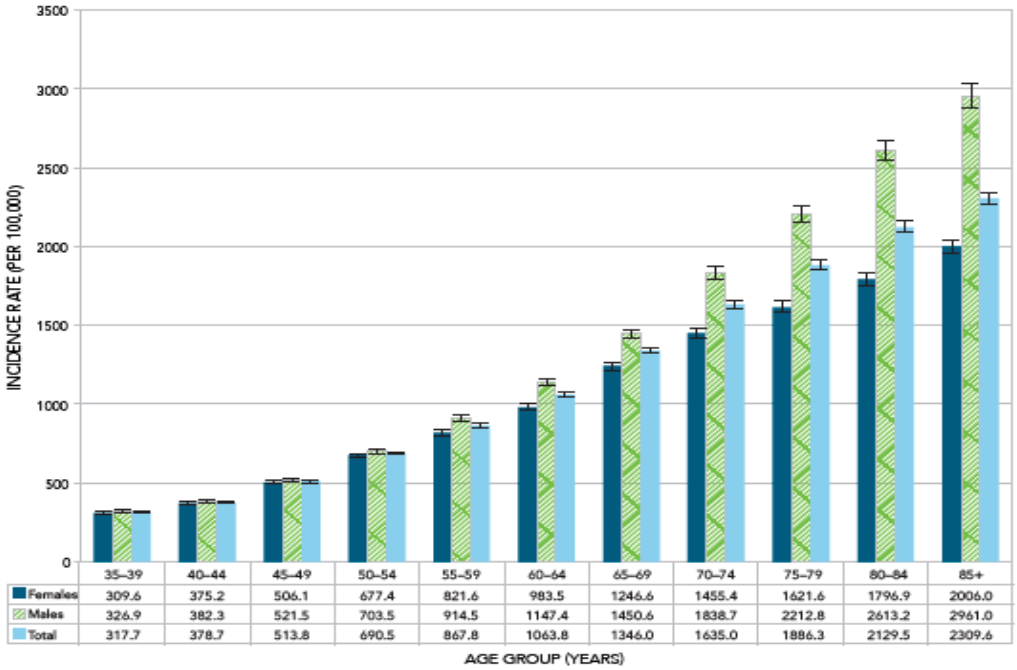


Figure 2. 12. Incidence rates of diagnosed COPD among Canadians aged 35 years and older, by age group and sex, Canada, 2011–2012

Source: Public Health Agency of Canada (2018)[7]

2.3.3. Saskatchewan

In 2011, Saskatchewan (SK) recorded an age-standardized COPD prevalence of 9%, the second lowest among the four western provinces after British Columbia (BC) and slightly lower than the national estimate of 9.4% [7]. Between 2001/02 and 2010/11, Saskatchewan had an increase in COPD prevalent cases from 1901 per year to more than 54,000 cases in 2010/11 [95]. Between 2001/02 and 2010/11, the age-standardized incidence rate of COPD fluctuated between 9 per 1000 and 10.3 per 1000 [96]. Over the 10 years, the COPD incidence rate was higher in males than in females [96]. In the 2014/2015 fiscal year, about 1 in 10 SK residents had COPD, which was more prevalent among residents aged 80 years and above, and 10% higher in males than females [97]. These observations are consistent with national and other provincial estimates [7].

In 2014/15, about one in every ten Saskatchewan residents (all ages) suffered from asthma [98]. The prevalence was about 5% higher in females compared to males [98]. SK had the second lowest age-standardized asthma prevalence of 10.2% in the western region of Canada in 2011/12 [7]. Although this value is slightly less than the national estimate of 10.8%, it is still

relatively high when compared to other provinces such as Newfoundland and Labrador (NL), New Brunswick (NB), BC, and Alberta (AB). Between 2001/02 and 2010/11, the net increase in asthma prevalence in SK was 3969 cases per year to more than 113,000 cases in 2010/11, which represents a prevalence of 10.3% in 2010/11.

Between 2010/11 and 2014/15, people with asthma in SK had 1.4 times higher age-adjusted risk of death and spent 1.7 more days in hospital compared to people without asthma [98]. During the same period, people with COPD had a 2.6 times higher risk of death and spent three times more days in the hospital compared to those without COPD [97]. These estimates are likely to increase with the ageing Canadian population, given that COPD incidence and prevalence increase with age [7,99]. This may likely lead to more pressure on the health care system.

2.4. Risk Factors for CRC

CRC are the result of the interaction between genetic and environmental factors over the life course, and for this reason, are referred to as "complex diseases" [12]. Genetically, a disease can either be inherited from the mother or the father or both parents through chromosomes [12]. Thus, a child is likely to suffer from CRC if any or both parents ever suffered from the disease. Several genes have been identified that influence susceptibility to CRC in both children and adults [100,101]. However, since altering the genes that contribute to these diseases is impracticable, at least for now, attention has been shifted to managing the environmental factors, also referred to as modifiable risk factors, that contribute to CRC.

Particulate exposures are the most extensively researched, potentially modifiable environmental risk factors for the development of CRC and include tobacco smoking and indoor/outdoor air pollutants[102–106]. Tobacco smoking, either through personal smoking or second-hand smoking, increases the risk of developing lung cancer, COPD and asthma [3] and is the most important risk in high-income countries [104]. While tobacco smoke is considered the primary cause of CRC, only a minority of smokers develop CRC [107,108], and non-smokers also develop CRC [109,110]. Exposure to outdoor air pollution has also been associated with incident COPD but is thought to contribute only modestly in comparison to tobacco smoke [111]. Globally and in low- and middle-income countries, exposure to biomass smoke is a major cause of CRC, with women and children suffering the most [104,112]. The contribution of the workplace environment to CRC is mostly through the inhalation of specific particles, gases, fumes or

smoke, but it can also be enhanced by smoking and pre-existing allergies [113]. It is estimated that occupational agents account for approximately 17% of adult asthma cases and 15-20% of COPD cases[113]. After adjusting for smoking in a case-control study, Govender et al. found high exposure to biological dust exposure, high cumulative mineral dust exposure, and high cumulative gas and fumes exposure, as well as the number of years of exposure to be significantly associated with COPD [114]. These authors also found risks of the above factors to be higher among smokers [114], suggesting that smoking may also heighten the effect of these risk factors.

Evidence is accumulating for the role of other exposures in the development of CRC, including those in utero, respiratory infections, nutritional status, and obesity [115–118]. Over the last decade, the role of socioeconomic status (SES) as a determinant of CRC has received increasing attention in the epidemiological literature [119,120]. SES is a complex, multidimensional construct representing an individual's or group's standing in society in terms of *power, prestige, and access to resources* [37]. The most common indicators of SES in Western society are income, education, and occupation, although many others exist.[53,58] A considerable body of research has accumulated reporting inverse, graded associations between SES and a variety of health status indicators [121–124], including CRC [19,30,125]. In the sections below, the conceptual underpinnings of SES as a determinant of health are presented, followed by an application to CRC.

2.4.1. SES

The Commission on Social Determinants of Health (CDSS) framework (Figure 2.12) which will be used to inform this study, draws on numerous theoretical perspectives to understand how health inequities arise and are maintained. Health inequities refer to *"health differences that are socially produced, systematic in their distribution across the population, and unfair"*[37,126] The first perspective, social causation, purports that various contextual factors determine health [127]. As shown in Figure 2.13, social, economic, and political contexts lead to the designation of life-defining socioeconomic statuses (e.g., educational attainment, income, occupation), resulting in contact with adverse/protective intermediary determinants, eventually translating into inequities in population health [37]. Intermediary determinants can be classified into material, psychosocial, and behavioural factors [37]. Material factors are associated with conditions of economic hardship and health-damaging conditions in the physical environment,

including access to healthy housing and proper nutrition; psychosocial factors include stressors, lack of social support, and stressful living conditions; and behavioural factors include smoking, diet and physical exercise [37]. Social causation perspectives highlight two causal mechanisms to explain health inequities: 1) differential exposure, whereby lower SES individuals are more likely to encounter adverse material, psychosocial and/or behavioral exposures, which in turn, increase the probability of health problems; and/or 2) differential susceptibility, whereby the effect of a potentially adverse exposure on health may be exacerbated when also exposed to low SES. SES can impact health at different levels of influence [128,129], and throughout the life course [128,130].

Also shown in Figure 2.13 is the second perspective, social selection, which posits that health determines SES rather than the vice versa. This phenomenon occurs through health exerting a strong effect on an individual's SES by hindering employment opportunities and subsequently income [131]. Thus, an individual's health can affect their attainment of SES, resulting in a social mobility pattern through which unhealthy individuals drift down the SES hierarchy and the healthy move up [37].

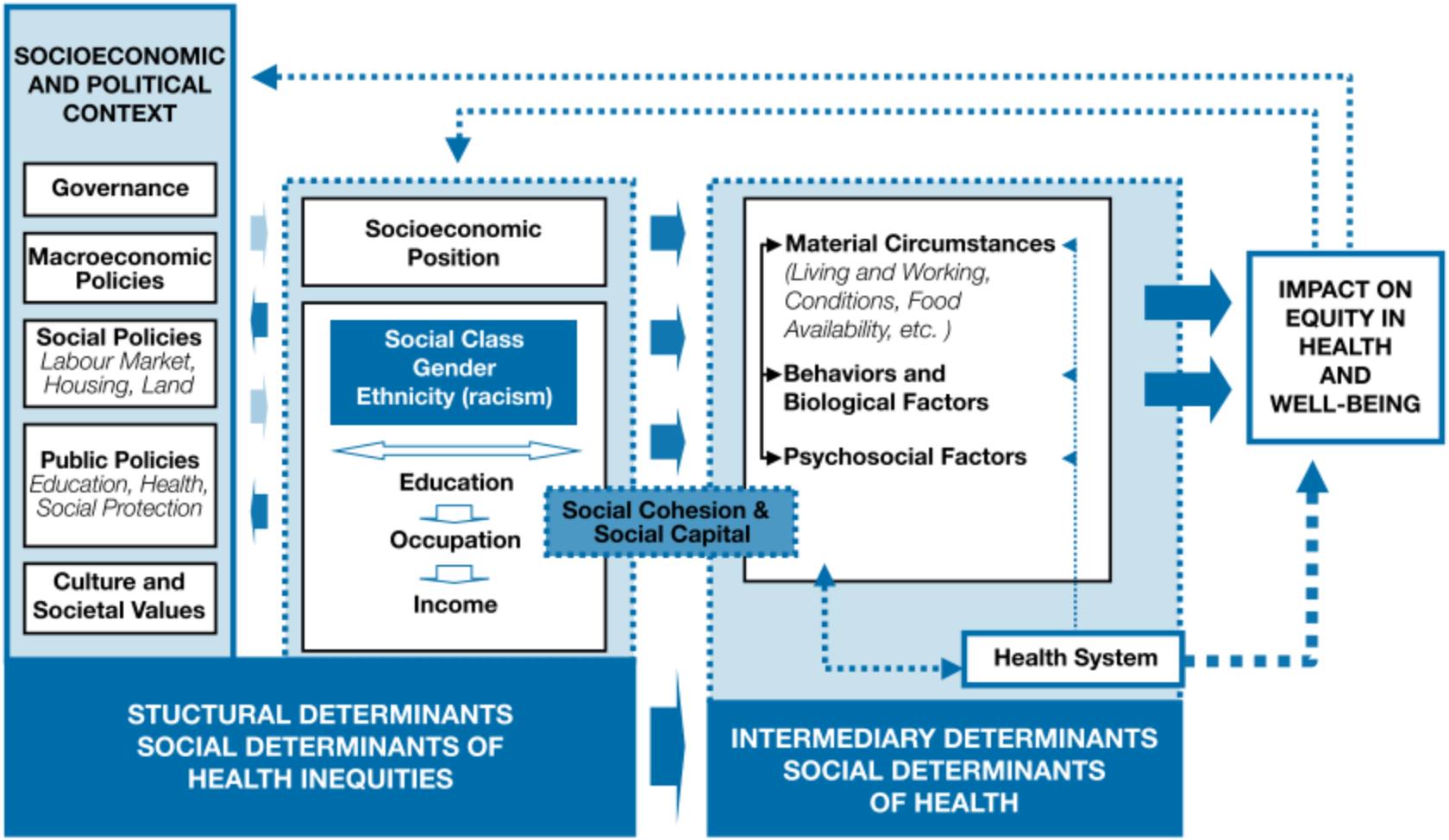


Figure 2. 13. Commission on the Social Determinants of Health Conceptual Framework

Source: Solar and Irwin (2010)[37]

2.6. SES and CRC

Numerous reviews have been published in the last decade that have concluded that SES is related to CRC [19,125,132]. One of the most sound methodologically was a systematic review of SES and COPD conducted in 2012 by Gerson et al [19]. After reviewing studies of higher quality published between 1996 and 2011, these authors concluded that substantial evidence existed to suggest lower SES is associated with higher COPD prevalence, mortality, and hospitalization, and that these relationships were robust, *given that they were consistent across gender, age, and population and appeared uninfluenced by how COPD was diagnosed, the SES measure used, or the disease outcome studied*. More recent reviews have drawn similar conclusions, including those focused on asthma [132].

The most common mechanism invoked to explain the SES patterning of CRC is differential exposure; that is, people lower in the SES hierarchy are more likely to be exposed to established risk factors for CRC, such as smoking and air pollution, which in turn lead to their higher risk of disease. Lower SES individuals are less likely to have health insurance or receive new drugs and technologies [133,134]. Research has shown that smoking prevalence continues to be higher among socially disadvantaged groups [135,136]. In addition, studies have shown higher rates of behavioural factors such as smoking, obesity, and substance use among low-income individuals [134,137,138]. Kim and Knesebeck conducted a meta-analysis and found lower income to be associated with subsequent obesity [122]. Individuals in financially strained families were also found to be more likely to smoke compared to those with sufficient emergency funds [139]. Education can promote health by reducing exposures to environmental risk factors such as smoking and second-hand smoking. Assari and Bazargan found that higher education in the United States was associated with a lower odds of exposure to second-hand smoking in the workplace [140]. Higher education was also found to be negatively associated with obesity and smoking, and high class occupation was found to be associated with less odds of smoking [141,142]. These behaviours are heavily influenced by challenging home environments and the community environment in which individuals live [143–145].

As noted in the most recent iteration of GOLD recommendations [146], the predominance of cross-sectional study designs in research examining the relationship between various risk factors and COPD poses interpretive challenges. The same holds true for studies of SES and CRC, with

the vast majority focused on SES predictors of prevalent SES, including those appearing in the reviews described above.

A search of the more recent scholarly literature was done to identify studies of a similar nature not included in previous reviews. Research examining the relationship between CRC and SES were searched March 2021 to July 2021 using Medline, PsycINFO, and Google Scholar, and included articles from 2012 to 2021. Search terms of keywords including "chronic respiratory conditions", "chronic respiratory diseases", "chronic obstructive pulmonary disease", "emphysema", "chronic bronchitis", "asthma", "socioeconomic status", "socioeconomic positions" were combined in exploring the literature. The search was limited to English language, adult human, and non-systematic review studies. The identified literature and key findings are presented in Table 2.

Out of the ten studies identified, seven examined prevalence of CRC [26,27,29,30,147–149] and three examined incidence [150–152]; six were focused on COPD [26,29,147–149,152], two on asthma [30,151], one on chronic bronchitis [150], and one a combination of COPD, bronchitis and emphysema [27]. The majority of studies examined household income, education, and occupations as indicators of SES, with a minority using other SES indicators such as area level income/deprivation, financial strain, and distance to a health centre [147,149,151,152]. China was the country of origin for three studies [26,147,148], Europe for four [29,30,149,152], with the remaining originating in the United States and Canada [27,150,151].

Similar to previous reviews, all of the studies reported statistically significant negative associations between SES and CRC, including incident CRC, with some caveats. Not all SES indicators were associated with CRC in every study, but it is difficult to identify definite patterns given the relatively low number of studies considered. The varied inclusion of particular covariates across studies also makes it more difficult to compare findings. Several studies pointed to more complex associations between SES and CRC, with effects dependent upon type of asthma, smoking status, and sex/gender. An important finding of this review is that relatively few studies have examined SES in relation to incident CRC, particularly in North America; analytical cross-sectional studies of prevalent disease make it challenging to establish the temporal patterning of exposure and disease. Further, although studies in rural settings were identified, the majority of these were focused on prevalence and set in China, calling into

question the transferability of study findings, given the different economic, political, and cultural context.

2.7. Summary

CRC are a major cause of suffering and death worldwide. Considerable research over the last decade demonstrates that lower SES is associated with CRC, but the nature of that association requires further clarification, particularly in a rural Canadian context. To that end, the present study uses longitudinal data from rural communities in a Canadian province to assess the relationship between SES and incidence of CRC.

Table 2. 1. List and summary of reviewed articles examining association between SES and CRC in adults

Author(s) year, country, sample	Study design	CRC	SES measure(s)	Confounders	Main results	Comments
Zhang et al. (2021)[26] China n=2,421; 40+ years old	Cross-sectional survey	Prevalent COPD Spirometry, respiratory symptoms, and risk factors	Education Occupation Household income	Age, residence, outdoor and indoor air pollution, cigarette smoking	Education: ≥13 yrs (Ref: <13) Male: **OR=0.27 (0.08 - 0.90)** Female: OR=1.94 (0.43 – 8.79) Occupation: White collar Male: **OR=0.62 (0.43 - 0.90)** Female: OR=0.60 (0.31 - 1.16) Income: Upper Male: **OR=0.63 (0.42 - 0.93)** Female: OR=0.82 (0.41 - 1.64).	Analyses stratified by sex: SES associated for men but not women
Cai et al. (2020)[147] China (rural) n=7,534; 35+ years old .	Cross-sectional	Prevalent COPD Physician diagnosed based on a spirometry	Education Household income Access to nearest healthcare center (good, less than 30 min walk; and poor (more than 30 min walk)	Age, current smoking status, outdoor exposure, indoor exposure to biomass fuel, tobacco cultivator status, weight, family history of lung disease	Education: High **OR = 0.89 (0.75 – 0.99)** Income: High OR= 1.02 (0.89 – 1.17) Access to medical service: Good **OR = 0.85 (0.73 – 0.99**)	Income not associated with COPD in this rural setting in China

Author(s) year, country, sample	Study design	CRC	SES measure(s)	Confounders	Main results	Comments
Ejlskov et al. (2018)[152] Denmark Birth cohort of 793,674 individuals born between 1961 and 1971	Retrospective cohort (birth cohort)	Incident COPD Diagnosis of COPD in Cause of Death register or Danish National Prescription Registry	Parental education Parental occupation Parental household income Composite SES (parental education, occupation, income)	Birth year and ethnicity (Danish or other)	Education (Reference High) Female Middle **HR = 1.62 (1.52 – 1.69)** Low **HR=2.30 (2.20 – 2.41)** Male Middle **HR = 1.67 (95% CI: 1.59 – 1.76)** Low **HR=2.14 (95% CI: 2.05 – 2.25)** Income (Reference High) Female Middle **HR = 1.43 (95% CI: 1.38 – 1.48)** Low **HR = 1.23 (95% CI: 1.19 – 1.28)** Male Middle **HR = 1.34 (95% CI: 1.29 – 1.39)** Low **HR = 1.30 (95% CI: 1.25 – 1.35)** Occupation (Reference High) Female Middle **HR = 1.76 (95% CI: 1.68 – 1.83)** Low **HR = 1.65 (95% CI: 1.59 – 1.70)** Male Middle	Focus on early life SES. Analyses stratified by sex and showed similar associations Strongest effect for education A significant portion of the early life SES effect on COPD was mediated through adult SES; however, some direct effects from early life SES on COPD remained

Author(s) year, country, sample	Study design	CRC	SES measure(s)	Confounders	Main results	Comments
					HR = 1.75 (95% CI: 1.68 – 1.81) Low **HR = 1.69 (95% CI: 1.63 – 1.74)** Composite index Female Middle **HR = 1.98 (95% CI: 1.89 – 2.09)** Low **HR = 2.64 (95% CI: 2.48 – 2.81)** Male Middle **HR = 1.98 (95% CI: 1.88 – 2.09)** Low **HR = 2.44 (95% CI: 2.30 – 2.59)**	
Coogan et al. (2016) [151] USA n=47,779 African American women; 21-69 years of age	Prospective cohort	Incident asthma Self-report of doctor-diagnosed asthma with concurrent use of asthma medication	Area level (neighborhood) composite SES measure (5 quintiles) Education (individual)	BMI, smoking status, alcohol consumption, second-hand smoking, abuse, racism, particulate matter, individual SES, health insurance	Neighborhood (Q5 – highest, ref category) Q1 HR=1.06 (0.89 – 1.26) Q2 HR=1.07 (0.90 – 1.28) Q3 **HR=1.21 (1.02 – 1.43)** Q4 HR=1.08 (0.91 – 1.28) Education (Reference High) Middle **HR = 1.13 (1.00-1.28)** Low **HR= 1.23 (1.05-1.44)**	Authors conclude that individual SES (education) appears to be more strongly associated with incident asthma than neighborhood SES

Author(s) year, country, sample	Study design	CRC	SES measure(s)	Confounders	Main results	Comments
Yin Peng et al. (2011)[148] China n=49,363; 15-69 years of age	Cross-sectional	Prevalent COPD Self-reported physician-diagnosed	Education Household income	Urban/rural, smoking status, age, sex, geographical location, passive smoking	**Urban** Education (Ref: High) Middle OR= 1.23 (0.93 – 1.57) Low **OR= 1.63 (1.32 – 2.13)** Income (Ref: High) Middle OR= 1.15 (0.91 – 1.45) Low **OR= 1.64 (1.28 – 2.09)** **Rural** Education (Ref: High) Middle OR= 1.21 (0.9 – 1.61) Low **OR= 1.74 (1.34 – 2.30)** Income (Ref: High) Middle OR= 0.84 (95% CI: 0.64 – 1.04) Low OR= 1.03 (0.84 – 1.26)	
Kainu et al. (2013)[29] Finland (urban) n=628; 20-79 years of age	Cross-sectional	Prevalent COPD Diagnosed using spirometry (GOLD criteria) and lower limit of normal (LLN)	Occupation	Sex/gender, age, smoking, occupation type	GOLD criteria Occupation (Ref: Professional) Manual work in industry OR = 2.30 (0.63 – 8.32) Manual work in service OR = 1.48 (0.48 – 5.09) Non-manual assistant employee **OR = 3.45 (1.21–9.84)** Other OR = 2.82 (0.70 – 11.4) LLN criteria Occupation (Ref: Professional) Manual work in industry OR = 3.23 (0.90–11.59) Manual work in service OR = 3.02 (0.93–9.82) Non-manual assistant employee **OR=4.44 (1.49–13.25)** Other **OR=5.58 (1.57–19.83)**	

Author(s) year, country, sample	Study design	CRC	SES measure(s)	Confounders	Main results	Comments
Schyllert et al. (2020)[30] Sweden n=6,854; 20-79 years of age	Cross-sectional	Prevalent asthma Affirmative answer to current physician diagnosis of asthma and symptoms or use of asthma medication in last year	Occupation Education	Age, smoking habits, body mass index categories, housing dampness and family history of asthma	Occupation (Ref: Professional and Executive) Non-manual (intermediate) OR = 1.21 (0.88 – 1.67) Non-manual (lower) OR = 1.12 (0.78 – 1.61) Manual worker in industry OR = 1.12 (0.80 – 1.56) Manual Worker in service **OR = 1.41 (1.03 – 1.91)** Education (Ref: Advanced tertiary) Basic OR = 1.10 (0.80 – 1.50) Upper Secondary OR = 1.02 (0.82 – 1.26) Compulsory School OR = 0.99 (0.74 – 1.35) Income (Ref: High income) Medium-high OR = 1.20 (0.93 – 1.55) Medium OR = 1.04 (0.80 – 1.35) Medium-low OR = 1.08 (0.83 – 1.40) Low OR = 1.16 (0.89 – 1.49)	The relationship between SES and asthma is complex Low occupation status and income more strongly associated with allergic asthma than with nonallergic asthma. Conversely, low education was associated with nonallergic asthma, whereas high education was associated with allergic asthma Interaction analysis indicated that women with low income (but not men) were at greater risk of current asthma
Levin et al. (2020)[149] Scotland (urban) n = 6,235, 16+ years of age	Primary care/mortality records, cross-sectional survey	Prevalent COPD Diagnosed using spirometry	Area level deprivation: Scottish Indicator for Multiple Deprivation (SIMD) 1 – 5 1 – Most deprived 5 – Least deprived	Sex, age	COPD age/sex adjusted prevalence per 1,000 (95% CI) SIMD 1 = 50.55 (49.82, 51.27) SIMD 2 = 34.26 (33.41, 35.11) SIMD 3 = 24.80 (24.00, 25.60) SIMD 4 = 16.56 (15.86, 17.26) SIMD 5 = 11.19 (10.69, 11.69)	Relationship between deprivation and COPD also dependent upon smoking status: • COPD decreased with increasing affluence for those not current smoking and also for those who had never smoked, with less of a gradient for the latter • The relationship between deprivation and COPD was less pronounced for current smokers compared with current non-smokers

Author(s) year, country, sample	Study design	CRC	SES measure(s)	Confounders	Main results	Comments
Janzen et al (2015)[27] Canada (rural) n = 8,261; 18+ years of age	Cross-sectional	Prevalent CRC (COPD, emphysema, bronchitis) Self reported doctor diagnosed	Household income Education Financial strain Occupation	Marital status, location within province, farm/nonfarm residence, urban accessibility	Financial strain (Ref: some money) Just enough money OR = 1.07 (0.84 – 1.37) Not enough money **OR = 1.48 (1.19 – 1.83)**	Relationship between education/income dependent upon age and sex • Low income associated with higher odds of CRC in younger but not older adults • Low education associated more strongly with CRC for younger women than younger men, but unrelated among older adults
Pahwa et al (2019)[150] Canada (rural) N=4,365; 18+ years of age	Prospective cohort	Incident chronic bronchitis Self reported doctor diagnosed	Household income Education	Age, sex, location within province, farm/nonfarm residence, family history of lung disease, ever asthma, ever allergy, BMI, household smoking, sex by age, household smoking by ever allergy	Household income (Ref: Highest) Lowest HR = 1.15 (0.79 – 1.74) Lower middle **HR = 1.98 (1.25– 3.13)** Upper middle HR = 1.65 (0.73 – 3.74)	Lower income but not lower education associated with incidence of chronic bronchitis

CHAPTER 3
METHODOLOGY

3.1. Introduction

This chapter describes the data source, variables, and statistical procedures of the study.

3.2. Data Source and Participants

Saskatchewan Rural Health Study (SRHS) adult data were used for this study [153,154]. Informed by a population health framework [155], the SRHS was a longitudinal study conducted in rural areas in Saskatchewan, Canada, to explore the hypothesis that individual and contextual factors are associated with respiratory health-related problems [153]. The questionnaire included assessment of participants' sociodemographic characteristics, health status, and respiratory-related exposures. Rural dwellers were defined as those *living in towns and municipalities outside the commuting zone of larger urban centers with a population of 10,000 or more* [156]. The SRHS consisted of a baseline survey in 2010, with a follow-up survey in 2014, each consisting of two components, a self-administered questionnaire, and a clinical assessment. Detailed information on participant recruitment and questionnaire design can be found elsewhere [153,154]. Only data from the self-administered questionnaires were used in this study.

The target population included tax-paying households located in the RMs and small towns of rural Saskatchewan. A multi-stage sampling design consisting of purposeful (stage 1) and random (stage 2) sampling strategies was used. Of the 297 RMs and 145 small towns, 48 RM and 16 small towns were selected from four quadrants (southeast, southwest, northeast, and northwest) of Saskatchewan. A random sample of 36 RMs was selected from each quadrant (9 out of 12 for each). Local councils for 32 out of 36 RMs and 15 out of 16 small towns gave their consent to participate in the study on behalf of the residents and provided a mailing list. Using the Dillman method to recruit participants 18 years of age and older, all prospective participants received a series of mail notifications [153,157]. A key informant in each household was asked to provide household-level information and individual information about each adult in the household. The baseline surveys were sent to 11,004 eligible households, of which 4,624 (42%) responded, representing 8,261 individuals.

Follow-up data were collected four years after the baseline phase, with self-administered questionnaires sent to all previously participating households [154]. To encourage a high retention

rate for the follow-up study, researchers kept in touch with participants through regular local newsletters/newspapers [154]. At the follow-up, 2,797 households comprising 4,867 men and women completed and returned the questionnaire. Out of 4,867, 4,741 individuals participated at both time points, and 126 individuals participated for the first time in 2014.

3.2.1 Sample size for the present study

Figure 3.1 shows the selection of participants for this study. The baseline and follow-up surveys resulted in initial samples of 8,261 and 4,867 individuals, respectively. Participants were then excluded if they had existing CRC at baseline, did not provide disease information at baseline, or if they participated in only one survey. The final sample size was 4,051 individuals, out of which 284 developed CRC at follow-up. Therefore, analyses were performed using a total sample size of 4,051.

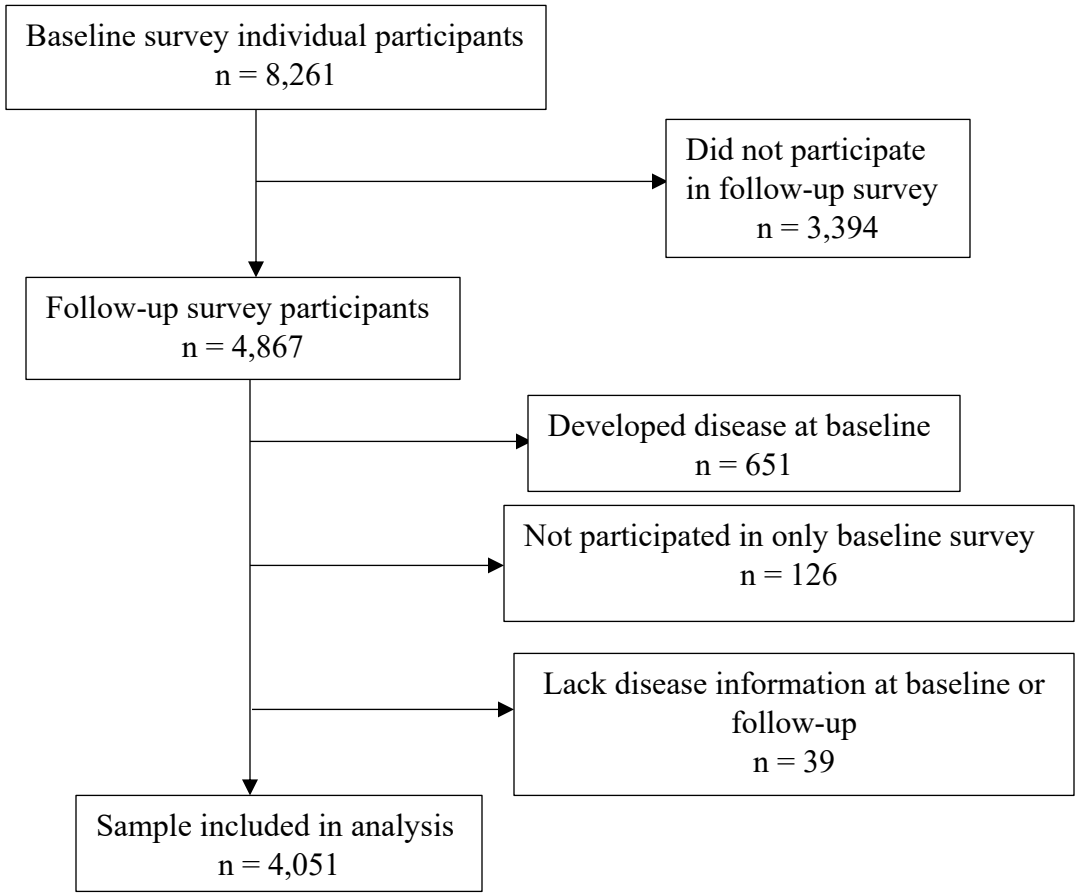


Figure 3. 1. Study population and sample size

3.3. Study Variables

3.3.1. Dependent variable

The primary outcome was chronic respiratory conditions (CRC), a derived dichotomous variable which consisted of the presence of one or more of the following: asthma, COPD, emphysema, or chronic bronchitis. Self-reported doctor-diagnosed asthma was examined with the questions *"Have you ever had asthma?"* and "*was it confirmed by a doctor?"* Participants had to answer 'yes' to both questions to be considered as an individual with asthma. Participants were also asked if a health professional ever said they had COPD, emphysema, or chronic bronchitis. An affirmative response to one or more of these questions, including asthma, indicated the presence of CRC. Combining these diseases into a single measure increased the number of new cases, and thereby, the statistical power of the analyses [158]. Only new cases that developed between 2010 and 2014 were of interest in this study.

3.3.2. Primary exposure variables

The primary independent variables of interest consisted of three SES indicators: educational level, household income, and financial strain. Educational level was assessed using the highest level of education of an individual and was categorized into three groups: (1) less than high school; (2) high school; and (3) postsecondary education. Two variables were considered for income: total household income and household income adequacy. Total household income was based on the response to the question *"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?"* Responses were grouped into four categories: (1) $29,999 or less; (2) $30,000 to $49,999; (3) $50,000 to $79,999; and (4) $80,000 or more. Household income adequacy, a derived variable combining total household income and the number of people living in the household, consisted of four categories: low income, low middle income, high middle income, and high income [153]. Financial strain was assessed using the question: *"At the end of the month, do you have some money left over, just enough, or not enough?"* [159]. Responses were categorized as some money, just enough money, and not enough money.

Demographic variables

Demographic variables included age, sex/gender, and marital status. Age was grouped into two categories, 18-64 years and 65 years and older [27]. Marital status was categorized as partnered (married and common law/living together) and not partnered (widowed, divorced/separated, and single/never married).

Geographical location

Geographical variables included quadrant location within the province (northwest; northeast; southwest; southeast), home location and degree of accessibility to urban areas and distance travelled for routine and ongoing medical care. Home location was examined using the question "*Where is your home located?*" Three responses, farm, in town, and acreage, were provided for this question; these responses were recategorized farm and non-farm (town and acreage). Metropolitan Influence Zone (MIZ) was used to assess degree of accessibility to urban areas (moderate, weak, no MIZ) [160]. Responses to the question, "*How far do you travel to receive routine and ongoing medical care?"* were categorized into the following: (1) 0 to 2 km; (2) less than 2km to 20km; (3) less than 20km to 51km; and (4) more than 51km.

Personal/behavioural

Physical activity was examined using the questions *"Do you exercise?"* and *"How long do you exercise?"* Responses were combined to create categories of no activity, less than 15 minutes, 15 – 30 minutes, 31 – 60 minutes, and more than 60 minutes. Smoking status (current smoker, ex-smoker, and never-smoker) was created using the questions *"Have you ever smoked cigarette?"* and *"Do you now smoke cigarettes?"* Alcohol consumption was examined with the question, *"During the past 12 months, how often did you drink alcoholic beverages?* This was grouped into four categories: (1) four or more times a week; (2) one to three times a week; (3) two to three times a month; and (4) one time or less than once a month. BMI was a derived variable computed from self-reported height and weight and grouped into *normal weight*, *overweight* and *obese*.

Environmental and occupational exposures

Occupational exposures were examined with the question *"Have you ever been exposed to any of the following in the workplace?"*; exposures considered were grain dust, mine dust, asbestos dust, wood dust, smoke from stubble burning, diesel fumes, welding fumes, solvent fumes, oil/gas well fumes, herbicides (to kill plants), fungicides (to treat grain), insecticides (to kill insects), molds, and radiation. Exposure to herbicides, insecticides and fungicides were combined into a single variable indicating pesticide exposure (yes/no).

Early life exposures

Early life exposures include whether respondents lived on a farm during their first year of life (*"Did you live on a farm during your first year of life?"*) and whether mother smoked while pregnant ("*Did your mother smoke while she was pregnant with you?")*

Co-morbidity, allergies and family history

Diabetes, heart disease, and heart attack were included as comorbidities and measured with a question of whether a doctor or primary care giver ever said the respondent had the disease. Also, allergy to various conditions (i.e., dog, cat, dust, household mold, grass, and pollen) were assessed; due to missing observations, a derived variable, any allergy (yes/no) was formed. Family history (mother or father) of lung disease (i.e., asthma, emphysema, or chronic bronchitis) was a dichotomous variable (yes/no).

Household living conditions

A household crowding index was derived by dividing participant responses to two questions: "*How many people live in your home?*" (numerator) and "*How many bedrooms do you have?*" (denominator), with three categories formed: (1) <1 person per bedroom; (2) 1 – 2 persons per bedroom; and (3) > 2 persons per bedroom [161]. Dampness was assessed with the question "*During the past 12 months, has there been water or dampness in your home from broken pipes, leaks, heavy rain, or floods?*" and mold with the question "*Does your home (including basement) frequently have a mildew odor or musty smell?*" The question "*Do any of the people who live in your house use any of the following products tobacco products?*" (cigarettes, cigars, pipes) was used to measure exposure to environmental tobacco smoke.

Natural gas as a source of fuel was assessed with the question "*Is natural gas primary fuel source to heat your home?*" The responses to all these questions were dichotomized into yes or no, with "don't know responses considered as missing.

3.4. Statistical Analysis

Stata version 15 was used for statistical analysis.

3.4.1. Descriptive analysis

The frequencies of baseline variables were presented in numbers and percentages. The incidence of CRC was calculated as the percentage of new CRC that occurred between 2010 and 2014 using the formula below.

$$Incidence\ of\ CRC = \frac{number\ of\ new\ case\ between\ 2010\ and\ 2014}{total\ valid\ number\ of\ particpants} \times 100,$$

where total valid participants include those who participated in both surveys and were disease-free at baseline. Incidences of the specific conditions which comprise CRC were also calculated, as well as for co-occurring COPD and asthma (ACOS). The incidence of CRC was further stratified by SES and all other variables used in the study. Cox proportional hazard regression analysis, described in more detail below, was also conducted to provide crude estimates (hazard ratios) of the association between CRC and each study exposure.

3.4.2. Multivariable modelling

The proportional Cox hazard model was used to estimate initially unadjusted and then adjusted hazard ratios for the study exposures. The time of origin was the baseline survey, and the follow-up survey was used to determine incident cases. Participants who did not develop CRC during the follow-up survey were considered as censored cases. Variables from the baseline survey were used in the Cox proportional hazards model. The proportional hazard model assumes that the hazard for an individual is a product of a baseline hazard (h_o) and an exponential function of a series of explanatory variables [162]. The hazard function can be interpreted as the risk of event at time t and can be expressed as,

$$h(t) = h_o(t)e^{\sum_{i=1}^{P} \beta_i X_i} \qquad (3.1)$$

where,

- *t* represents the survival time.
- *h(t)* is the hazard function determined by a set of *i* covariates.
- β_i are the coefficients estimates and they measure the impact of covariates.
- $h_0(t)$ is the baseline hazard. It corresponds to the value of the hazard if all covariates are zero.

The model can also be written as a multiple linear regression of the logarithm of the hazard on the variable X_i, with the baseline hazard being an intercept term that varies with time as represented in Equation 2,

$$lnh(t) = lnh_o(t) + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_p X_p \qquad (3.2)$$

Time is one of the primary assumptions of the Cox proportional hazard model [163]. The assumption of the proportional hazard model is tested using the Schoenfeld residual [164]. The statistical form of this test was chosen over the graphical form for convenience. Goodness of fit for the final model was assessed using the Cox-Snell residuals and Nelson-Aalen cumulative hazard function [165], as well as the Harrell's C statistics [166].

3.4.2.1. Model Building

Variables with p-values < 0.20 in bivariate analyses and those identified as clinically/ biologically important were chosen for multivariable analysis. All selected variables were entered into the model at once. With the exception of the two income variables, the SES variables were not highly correlated; therefore, all three SES indicators (but excluding total household income) were included together in the multivariable analysis. Two-way interactions between SES indicators, age, and sex/gender were assessed. Variables with p-values <0.05 were considered statistically significant and remained in the final model.

3.4.3. Missing data

Table 3.1 shows the number and proportion of missing values for each study exposure. Missing data can be dealt with in several way including listwise deletion and multiple imputation [167]. The listwise deletion approach deletes all observations that have missing data in any of the variables, which reduces the sample size used in estimation. Even though this approach is convenient and simple, it can reduce statistical power, produce bias estimates, and reduce efficiency [168,169] if data is missing not at random (MNAR) or missing at random (MAR). However, if data is missing completely at random (MCAR), the estimates will not be biased, but there will be reduced statistical power [170]. For each variable with more than 5% missing observations, a chi square test was used to determine whether there was a statistically significant association between missingness and CRC incidence; no statistically significant differences were present, hence, the list-wise deletion method was used (Table 3.2).

Table 3. 1. Number and percentage of missing observations for each study variable

Variables	Number	%
Pesticides	41	1.01
Household density	22	0.54
Alcohol consumption	10	0.25
Live on farm 1st year	36	0.89
Cigarette smoking at home	19	0.47
Household mold	103	2.54
Heart disease	46	1.14
Household dampness	31	0.77
Household natural gas use	16	0.39
Educational level	33	0.81
Financial Strain	370	9.13
Household income adequacy	573	14.14
BMI	178	4.39

Variables	Number	%
Home location	18	0.44
Smoking status	12	0.30
Household income	551	13.60
Exercise	137	3.38
Wood dust	41	1.01
Asbestos dust	41	1.01
Mine dust	41	1.01
Diabetes	20	0.49
Household molds	41	1.01
At least one parent had lung disease	483	11.92
Marital status	13	0.32
Heart disease	46	1.14

Table 3.2. Distribution of missing and non-missing observations for variables with more than 5% missing observations

Variables	CRC		
	Not developed (%)*	Developed (%)*	P-value
Household income adequacy			0.24
Not missing	2993 (86.33)	238 (83.80)	
Missing	474 (13.67)	46 (16.20)	
Financial strain			0.97
Not missing	3152 (90.91)	258 (90.85	
Missing	315 (9.09)	26 (9.15)	
At least one parent had lung disease			0.71
Not missing	3065 (88.40)	249 (87.68)	
Missing	402 (11.60)	35 (12.32)	

**Column percentages are in parenthesis*

3.4.4. Sensitivity analysis

The individual components of CRC (COPD and asthma) were also analysed separately to understand which of them may be driving any observed effects in the CRC model. In addition, ACOS was considered; that is, individuals who self-reported both COPD and asthma in 2014 [73]. Self-reported COPD and asthma is one of the definitions used to identify ACOS, albeit there is currently no universally accepted definition [69,171,172]. This analysis aimed to determine whether ACOS and CRC have different risk factors.

CHAPTER 4
RESULTS

4.1. Introduction

This chapter commences by presenting frequency distributions of baseline study variables. The incidence of CRC is then described, followed by univariable Cox proportional hazard model results. The multivariable Cox proportional hazard model results for CRC are described, followed by sensitivity analysis results.

4.2. Descriptive statistics

4.2.1. Baseline Characteristics of the study population

The baseline characteristics of the study population are presented in **Table 4. 2,** and shows that 14% of participants had CRC. A slightly higher proportion of women than men participated, and the vast majority (81%) were partnered. Just over two-thirds were less than 65 years old and 41% reported that their homes were located on farms. Nearly 60% were from the north study regions and most (82%) in weak or no MIZs. Nearly two-thirds of participants were overweight or obese. One-third of respondents reported having attained high school or less, just over one-half had some money left over at the end of each month, and nearly two-thirds were in the upper middle or high Household income adequacy groups. **Table 4.1** also shows the frequency distribution of all other variables in the study.

Table 4. 1. Study population characteristics at baseline

Variables	Number/Total	Percent (%)
CRC	1185/8156	14.53
COPD	669/8,151	8.21
Asthma	715/8,261	8.66
ACOS	199/8,256	2.41
Demographics		
Age group		
18-64 years	5,775/8,387	68.86
65+ years	2,612/8,387	31.14
Sex/gender		
Male	4,064/8,258	49.21
Female	4,194/8,258	50.79
Marital status		
Partnered	6,780/8,226	82.42
Not partnered	1,446/8,226	17.58
Geographical location		
Quadrant		

Variables	Number/Total	Percent (%)
Southwest	1,538/8257	18.63
Southeast	1,792/8257	21.70
Northeast	2,400/8257	29.07
Northwest	2,527/8257	30.60
Home location		
Farm	3,445/8208	41.97
Non-farm	4,763/8208	58.03
MIZ		
No MIZ	2,314/8257	28.02
Weak MIZ	4,573/8257	55.38
Moderate MIZ	1,370/8257	16.59
Distance to health care		
0-2 Km	1,530/5,158	29.66
>2-20	1,135/5,158	22.00
>20-51 Km	1,207/5,158	23.40
>51 Km	1,286/5,158	24.93
SES		
Educational level		
Less than high school	2,127/8159	26.07
High school	2,814/8159	34.49
Postsecondary	3,218/8159	39.44
Financial Strain		
Not enough money	4,427/7450	59.42
Just enough money	1,595/7450	21.41
Some money	1,428/7450	19.17
Household income adequacy		
Lower income	324/6977	4.64
Lower middle income	1,216/6977	17.43
Upper middle income	2,305/6977	33.04
High income	1,410/6977	44.89
Household conditions		
Crowding index		
< 1	5,575/8,387	66.47
1 – 2	2,586/8,387	30.83
> 2	226/8,387	2.69
Household natural gas use		
Yes	5,587/8228	67.90
No	2,641/8228	32.10
Household dampness		
Yes	1,566/8195	19.11
No	6,629/8195	80.89
Household Mold		
Yes	2,803/8,096	34.62
No	5,293/8,096	65.38
Cigarette smoking at home		
Yes	1,249/8,216	15.20
No	7,011/8,216	84.80
Workplace exposures		
Stubble smoke		
Yes	3,255/7685	42.36

Variables	Number/Total	Percent (%)
No	4,430/7685	57.64
Diesel fumes		
Yes	4,751/7,813	60.81
No	3,062/7,813	39.19
Welding fumes		
Yes	3,327/8,096	41.09
No	4,769/8,096	58.91
Solvent fumes		
Yes	2,865/8,096	35.39
No	5,231/8,096	64.61
Radiation		
Yes	688/8,096	8.50
No	7,408/8,096	91.50
Wood dust		
Yes	3,112/8,096	38.44
No	4,984/8,096	61.56
Grain dust		
Yes	5,523/8,096	68.22
No	2,573/8,096	31.78
Asbestos dust		
Yes	543/8,096	6.71
No	7,553/8,096	93.29
Mine dust		
Yes	445/8,096	5.50
No	7,651/8,096	94.50
Oil/gas fumes		
Yes	1,944/8,096	24.01
No	6,152/8,096	75.99
Pesticides		
Yes	4,517/8,096	55.79
No	3,579/8,096	44.21
Lifestyle		
Smoking status		
Current smoker	968/8,216	11.78
Ex-smoker	2,923/8,216	35.58
Never smoker	4,325/8,216	52.64
Alcohol consumption		
Never	1,472/8,225	17.90
1 or less a month	2,484/8,225	30.20
2-3times a month	1,448/8,225	17.60
1-3times a week	1,955/8,225	23.77
4 or more a week	866/8,225	10.53
Physical activity		
None	3,421/7,908	43.26
Less than 15 min	456/7,908	5.77
15-30 min	2,144/7,908	27.11
31-60 min	1,415/7,908	17.89
More than 60 min	472/7,908	5.97
BMI		
Normal	2,345/7,841	29.91

Variables	Number/Total	Percent (%)
Overweight	3,207/7,841	40.90
Obese	2,289/7,841	29.19
Early life exposures		
1st year of life on farm		
Yes	5,555/8,171	67.98
No	2,616/8,171	32.02
Mother smoke during pregnancy		
Yes	1,214/7,307	16.61
No	6,093/7,307	83.39
Allergies		
Any allergy		
Yes	2,670/8,261	32.32
No	5,591/8,261	67.68
Family history		
At least one parent had lung disease		
Yes	1,962/7,049	27.83
No	5,087/7,049	72.17
Morbidity		
Heart disease		
Yes	610/8,096	7.54
No	7,481/8,096	92.46
Diabetes		
Yes	759/8,154	9.31
No	7,395/8,154	90.69

CRC – Chronic respiratory conditions; COPD – Chronic obstructive pulmonary disease; ACOS – Asthma COPD overlap syndrome; MIZ – Metropolitan influence zones; SES – Socioeconomic status

4.2.2. Incidence

The crude incidence of CRC over the four years was 7.01%. Each SES indicator showed a statistically significant inverse, dose-response association with CRC, with the crude incidence increasing as SES was decreasing. (**Figure 4.1** and **Table 4.2**). Older residents had a significantly higher incidence of CRC (9.11%) than those younger (6.15%), whereas there was no difference by sex/gender. A higher proportion of farm-dwellers (7.16%) than town residents developed CRC (6.85%), as did those not partnered (9.23%) compared to partnered (6.70%). The remaining exposures were all statistically significantly associated with CRC incidence, with the exception of home location, MIZ, distance to healthcare, household crowding index, household dampness, household mold, all workplace exposure variables (except exposure to asbestos), alcohol consumption, physical activity, BMI, first year of life on farm, and mother smoked during pregnancy,

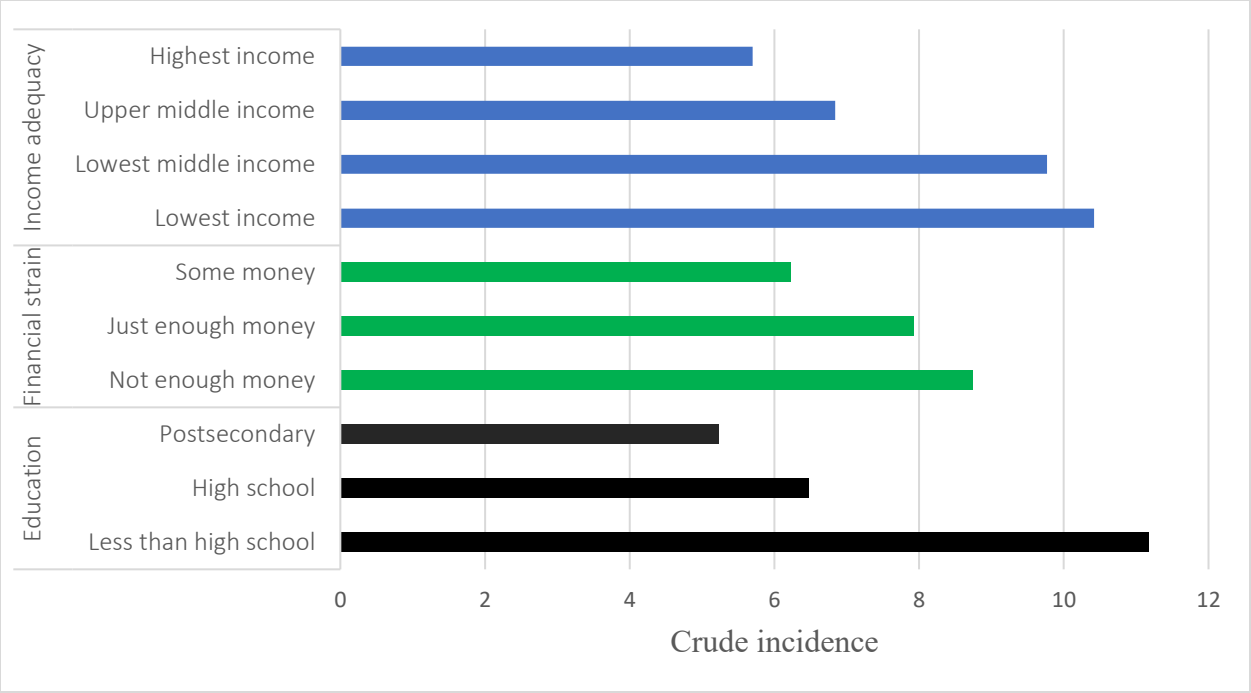


Figure 4. 1. Crude incidence of CRC stratified by SES indicators

Table 4. 2. Crude incidence of CRC by risk factors and unadjusted hazard ratios

Variables	Incidence of CRC		Unadjusted hazard ratio (HR)	
	Yes/total (%)	p-value	HR (95% CI)	p-value
Demographics				
Age group		<0.001	Ref	0.006
18-64 years	177/2,877 (6.15)		**1.55 (1.21 – 1.99)**	
65+ years	107/1,174 (9.11)			
Sex/gender		0.624		0.837
Male	136/1,977 (6.88)		Ref	
Female	148/2,074 (7.14)		1.02 (0.82 - 128)	
Marital status		0.001		0.165
Partnered	235/3,507 (6.70)		1.26 (0.91 – 1.74)	
Not partnered	49/531 (9.23)		Ref	
Geographical location				
Quadrants		0.016		0.018
Southwest	42/724 (5.80)		Ref	
Southeast	80/860 (9.30)		**1.73 (1.16 – 2.57)**	
Northeast	77/1,191 (6.47)		1.11 (0.74 – 1.66)	
Northwest	85/1,276 (6.66)		1.29 (0.87 – 1.92)	
Home location		<0.001		0.472
Farm	133/1,858 (7.16)		1.09 (0.86 – 1.40)	
Non-Farm	149/2,175 (6.85)		Ref	
MIZ		<0.001		0.066
No MIZ	72/1,108 (6.50)		Ref	
Weak MIZ	173/2,231(7.75)		1.27 (0.96 – 1.70)	

Variables	Incidence of CRC		Unadjusted hazard ratio (HR)	
	Yes/total (%)	p-value	HR (95% CI)	p-value
Moderate MIZ	39/712 (5.48)		0.88 (0.59 – 1.32)	
Distance to healthcare		<0.001		0.731
0-2 Km (=< Q1)	50/748 (6.68)		Ref	
>2-20 Km (Q1-Q2)	41/541 (7.58)		1.23 (0.81 – 1.88)	
>20-51 Km (Q2-Q3)	44/622 (7.07)		1.09 (0.72 – 1.67)	
>51 Km (>Q3)	48/608 (7.89)		1.22 (0.81 – 1.85)	
SES				
Educational level		<0.001		<0.001
Less than high school	105/939 (11.18)		**2.00 (1.51 – 2.66)**	
High school	88/1,358 (6.48)		1.18 (0.88 – 1.58	
Postsecondary	90/1,721 (5.23)		Ref	
Financial strain		0.031		0.199
Just enough money	60/757 (7.93)		1.19 (0.87 – 1.63)	
Not enough money	55/629 (8.74)		1.33 (0.95 – 1.86)	
Some money	143/2,295 (6.23)		Ref	
Household income		<0.001		0.002
<= $29,999	56/535 (10.47)		**1.94 (1.01 – 3.73**	
$30,000- $49,999	66/771 (8.56)		**1.90 (1.33 –2.71)**	
$50,000- $79,999	43/871 (4.94)		1.21 (0.89 – 1.66)	
>=$80,000	77/1,323 (5.87)		Ref	
Household income adequacy		<0.001		0.002
Lower income	10/96 (10.42)		**1.95 (1.36 – 2.79)**	
Lower middle income	51/522 (9.77)		**1.46 (1.03 – 2.08)**	
Upper middle income	84/1,228 (6.84)		0.84 (0.57 – 1.23)	
High income	93/1,632 (5.70)		Ref	
Household conditions				
Crowding index		0.096		0.178
< 1	194/2,822 (6.87)		Ref	
1 – 2	85/1,197 (7.10)		1.01 (0.77 – 1.33)	
> 2	5/32 (15.63)		2.08 (0.96 – 4.51)	
Molds		<0.001		0.005
Yes	118/1393 (8.47)		**1.41 (1.11 – 1.79)**	
No	162/2,617 (6.19)		Ref	
Natural gas		<0.001		0.032
Yes	170/2,683 (6.34)		Ref	
No	114/1,352 (8.43)		**1.31 (1.02 – 1.68)**	
dampness		0.475		0.851
Yes	56/768 (7.29)		Ref	
No	226/3,252 (6.95)		1.03 (0.76 – 1.39)	
Cigarette smoking at home		0.001		0.044
Yes	42/444 (9.46)		**1.40 (1.01 – 1.95)**	
No	241/3,588 (6.72)		Ref	
Workplace exposure				
Stubble smoke		<0.001		0.436
Yes	120/1,627 (7.38)		1.10 (0.86 – 1.41)	
No	147/2,217 (6.63)		Ref	
Diesel fumes		<0.001		0.558

Variables	Incidence of CRC		Unadjusted hazard ratio (HR)	
	Yes/total (%)	p-value	HR (95% CI)	p-value
Yes	169/2,377 (7.11)		1.08 (0.84 – 1.37)	
No	102/1,521 (6.71)		Ref	
Welding fumes		<0.001		0.596
Yes	110/1,650 (6.67)		Ref	
No	170/2,360 (7.20)		1.07 (0.84 – 1.35)	
Solvent fumes		<0.001		0.917
Yes	100/1,439 (6.95)		Ref	
No	180/2,571 (7.00)		1.01 (0.79 – 1.29)	
Radiation		<0.001		0.315
Yes	20/337 (5.93)		Ref	
No	260/3,673 (7.08)		1.27 (0.80 – 2.02)	
Wood dust		<0.001		0.062
Yes	126/1,593 (7.91)		1.25 (0.99 – 1.59)	
No	154/2,417 (6.37)		Ref	
Grain dust		<0.001		0.521
Yes	194/2,746 (7.06)		1.09 (0.85 – 1.39)	
No	86/1,264 (6.80)		Ref	
Asbestos dust		<0.001		0.039
Yes	28/255 (10.98)		**1.55 (1.02 – 2.35)**	
No	252/3,755 (6.71)		Ref	
Mine dust		<0.001		0.125
Yes	21/225 (9.33)		1.41 (0.91 – 2.18)	
No	259/3,785 (6.84)		Ref	
Oil/gas well fumes		<0.001		0.767
Yes	64/907 (7.06)		Ref	
No	216/3,103 (6.96)		1.04 (0.79 – 1.39)	
Pesticides		<0.001		0.322
Yes	166/2,302 (7.21)		1.13 (0.89 – 1.43)	
No	114/1,708 (6.67)		Ref	
Lifestyle				
Smoking status		<0.001		<0.001
Current smoker	37/367 (10.16)		**1.85 (1.27 – 2.69)**	
Ex-smoker	128/1,405 (9.11)		**1.74 (1.35 – 2.24)**	
Never smoker	119/2,270 (5.24)		Ref	
Alcohol consumption		0.054		0.435
never	55/628 (8.76)		Ref	
1 or less a month	90/1,201 (7.49)		0.93 (0.66 – 1.32)	
2-3times a month	44/750 (5.87)		0.78 (0.52 – 1.17)	
1-3times a week	58/1,004 (5.78)		0.75 (0.51 – 1.08)	
4 or more a week	36/458 (7.86)		0.98 (0.64 – 1.50)	
Physical activity		0.035		0.263
None	120/1,624 (7.39)		Ref	
Less than 15 min	22/217 (10.14)		1.22 (0.77 – 1.93)	
15-30 min	76/1,101 (6.90)		0.92 (0.69 – 1.23)	
31-60 min	37/748 (4.95)		0.69 (0.47 – 1.00)	
More than 60 min	14/224 (6.25)		0.97 (0.56 – 1.69)	
BMI		0.522		0.371
Normal	68/1,127 (6.03)		Ref	

Variables	Incidence of CRC		Unadjusted hazard ratio (HR)	
	Yes/total (%)	p-value	HR (95% CI)	p-value
Overweight	123/1,658 (7.42)		1.23 (0.92 – 1.64)	
Obese	80/ 1,088 (7.35)		1.14 (0.82 – 1.57)	
Early life exposures				
1st year of life on farm		0.481		0.114
Yes	208/2,843 (7.32)		1.24 (0.95 – 1.63)	
No	74/1,172 (6.31)		Ref	
Mother smoke		0.589		0.808
Yes	42/568 (7.39)		Ref	
No	215/3,026 (7.11)		1.04 (0.75 – 1.45)	
Allergies				
Any allergy		<0.001		<0.001
Yes	112/1,128 (9.93)		**1.72 (1.36 – 2.17)**	
No	172/2,923 (5.88)		Ref	
Family history				
One parent had lung disease		0.001		<0.001
Yes	85/906 (9.38)		**1.60 (1.23 – 2.06)**	
No	164/2,662 (6.16)		Ref	
Morbidity				
Heart disease		0.026		0.002
Yes	32/278 (11.51)		**1.78 (1.23 – 2.59)**	
No	247/ 3,727(6.63)		Ref	
Diabetes		0.021		0.017
Yes	33/331 (9.97)		**1.55 (1.08 – 2.22)**	
No	250/3,700 (6.76)		Ref	

4.3. Association between independent variables and CRC

4.3.1. Univariable analysis

The results of the univariable Cox proportional hazard regression analysis are also presented in Table 4.2. Older age was associated with an increased risk of CRC (HR = 1.55: 95% CI: 1.21 – 1.99). All the SES indicators showed an inverse graded association with CRC, but only education and household income adequacy were statistically significant. Specifically, having less than a high school education was associated with a 2 times higher risk (HR=2.0; 95% CI: 1.51 – 2.66) of developing CRC compared to those with post-secondary education. Compared with high income adequacy households, the risk of CRC was greater among those from low (HR=1.96; 95% CI: 1.01 – 3.73) and middle-income adequate households (HR=1.46; 95% CI: 1.33 – 2.71). Other statistically significant risk factors identified in the univariable analysis included the southeast quadrant compared to southwest (1.73; 95% CI: 1.16 – 2.57), the presence of mold (1.41; 95% CI: 1.11 – 1.79), not using natural gas in household (1.31; 95% CI: 1.02 – 1.68), exposure to household cigarette smoking (1.40; 95% CI: 1.01 – 1.95), exposure to

asbestos dust (1.55; 95% CI: 1.02 – 2.35), and being a current smoker (1.85; 95% CI: 1.27 – 2.69) or ex-smoker (1.74; 95% CI: 1.35 – 2.24) compared to a non-smoker. Other variables associated with CRC included any allergy (1.72; 95% CI: 1.36 – 2.17), having a parent with lung disease (1.60; 95% CI: 1.23 – 2.06), having heart disease (1.78; 95% CI: 1.23 – 2.59), and being diabetic (1.55; 95% CI: 1.08 – 2.22).

Other variables with p-value less than 0.20 and were included at multivariable modeling stage were age, marital status, MIZ, household crowding index, exposures to wood, asbestos, and mine dusts, and staying on farm in first year. Sex was also included due to its biological importance.

4.3.2. Multivariable analysis

The Schoenfeld residual test for the proportional hazard assumption showed a p-value of 0.24, suggesting that the assumption was not violated. The plotted Cox-Snell residual and Nelson-Aalen cumulative hazard (**Figure 4.1**) indicated that the model was fairly well-fitted. The Harrel's C (goodness of fit) was 0.69, indicating that survival time could be correctly ordered for pairs of respondents 69% of the time, based on the factors included in the model.

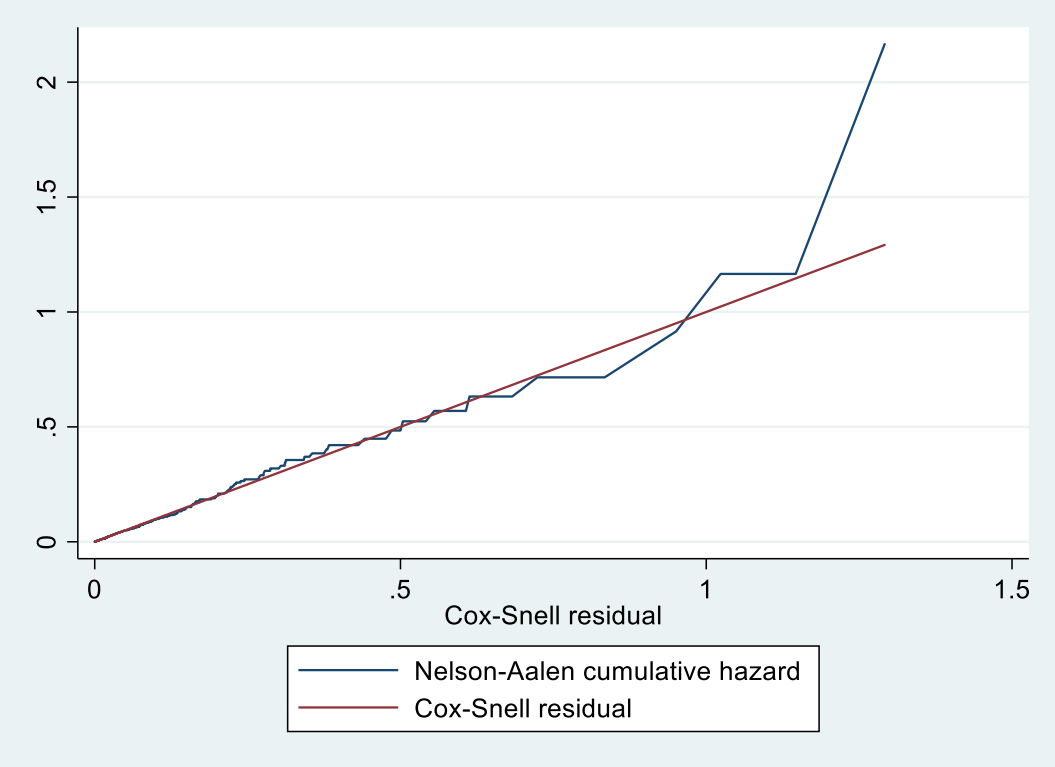


Figure 4. 2. Cox-Snell residual and Nelson Aalen cumulative hazard plot

4.3.2.1. Multivariable model results

After adjusting for demographic characteristics, household conditions, early life experience, workplace exposures, and co-morbidities, only household income adequacy was associated with the risk of CRC among the SES indicators (**Error! Reference source not f ound.**). The result of the final model is summarized in (**Error! Reference source not found.**). Compared with high income adequacy participants, those with lower income adequacy and lower middle income adequacy had 2.22 times (95% CI: 1.01 – 4.89) and 1.66 times (95% CI: 1.08 – 2.56) higher risk of CRC, respectively. Other factors associated with an increased risk of CRC were: not using natural gas as a source of heat (HR = 1.57; 95% CI: 1.15 – 2.15), being an ex-smoker compared to current smoker (HR = 1.65; 95% CI: 1.22 – 2.23), having at least one parent with lung disease (HR = 1.54; 95% CI: 1.14 – 2.07), having diabetes (HR = 1.64; 95% CI: 1.10 – 2.44) and having any allergy (HR = 1.66; 95% CI: 1.22 – 2.27). None of the interaction terms tested were statistically significant.

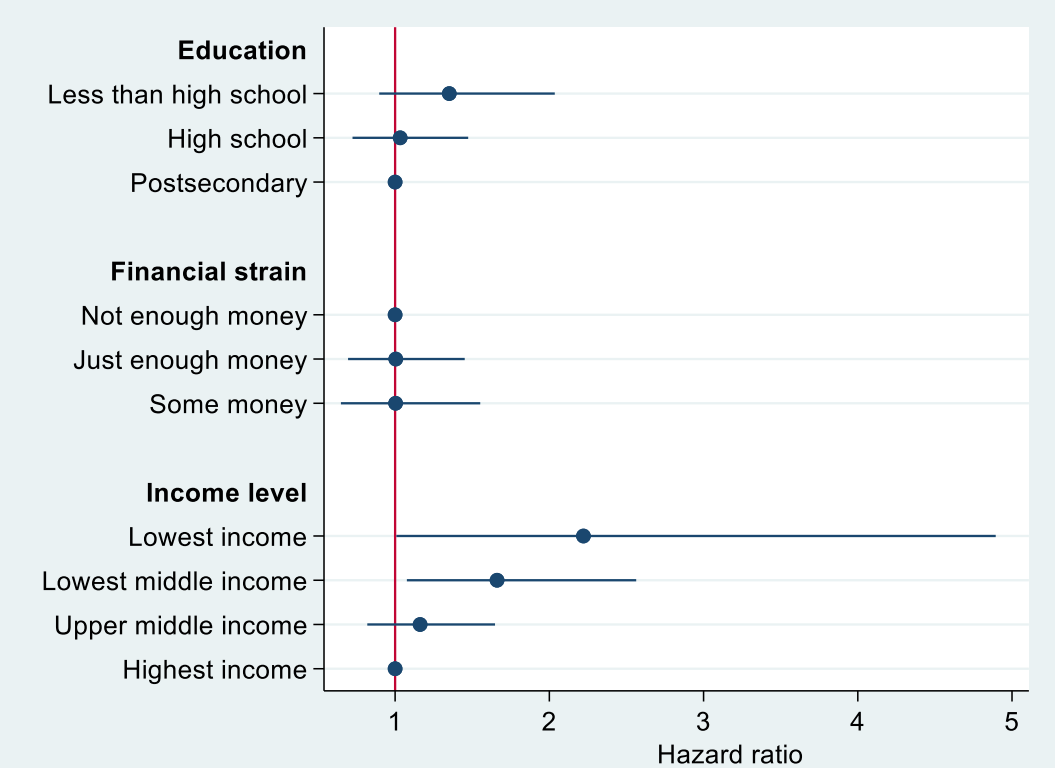


Figure 4. 3. Adjusted* hazard ratios plot of SES indicators on chronic respiratory conditions

**Adjusted for age, sex/gender, quadrant, MIZ, any allergies, heart disease, diabetes, at least one parent had lung disease, household mold, use of natural gas, household crowding index, cigarette smoking in household, exposures to wood, asbestos, and mind dusts, smoking status and BMI.*

Table 4. 3. Final multivariable Cox proportional hazard model results of the association between chronic respiratory conditions, SES indicators and other risk factors.

Variables	HR (95% CI)
Age group	
(18-64) years	1.00 (Ref)
65+ years	1.22 (0.87-1.72)
Sex	
Male	1.00 (Ref)
Female	1.21 (0.88-1.66)
Quadrants	
Southwest	1.00 (Ref)
Southeast	1.63 (0.98-2.73)
Northeast	1.05 (0.62-1.77)
Northwest	1.47 (0.86-2.51)
MIZ	
No MIZ	1.00 (Ref)
Weak MIZ	1.35 (0.93-1.95)

Variables	HR (95% CI)
Moderate MIZ	0.92 (0.51-1.67)
Education	
Less than high school	1.35 (0.90-2.04)
High school	1.03 (0.72-1.47)
Postsecondary	1.00 (Ref)
Financial Strain	
Not enough money	1.00 (0.65-1.55)
Just enough money	1.00 (0.69-1.45)
Some money	1.00 (Ref)
Household income adequacy	
Lower income	**2.22 (1.01-4.89)**
Lower middle income	**1.66 (1.08-2.56)**
Upper middle income	1.16 (0.82-1.65)
High income	1.00 (Ref)
Mold	
Yes	0.93 (0.67-1.29)
No	1.00 (Ref)
Natural gas	
Yes	**1.57 (1.15-2.15)**
No	1.00 (Ref)
Crowding index	
< 1	1.00 (Ref)
1 – 2	0.96 (0.69-1.32)
> 2	2.00 (0.72-5.58)
Cigarette smoking at home	
No	1.00 (Ref)
Yes	0.91 (0.54-1.52)
Wood dust	
No	1.00 (Ref)
Yes	1.08 (0.76-1.52)
Asbestos dust	
No	0.86 (0.49-1.50)
Yes	1.00 (Ref)
Mine dust	
No	0.63 (0.37-1.09)
Yes	1.00 (Ref)
Smoking status	
Current smoker	1.65 (0.90-3.04)
Ex-smoker	**1.68 (1.22-2.32)**
Non-smoker	1.00 (Ref)
BMI	
Normal (< 25)	1.00 (Ref)
Overweight (>=25 and <=30)	1.32 (0.94-1.88)
Obese (>30)	0.99 (0.67-1.48)
1st year of life on farm	
No	1.00 (Ref)
Yes	1.03 (0.75-1.43)
At least 1 parent had lung disease	
No	1.00 (Ref)
Yes	**1.54 (1.14-2.07)**

Variables	HR (95% CI)
Any allergy	
No	1.00 (Ref)
Yes	**1.66 (1.22-2.27)**
Heart disease	
No	1.00 (Ref)
Yes	1.11 (0.64-1.92)
Diabetes	
No	1.00 (Ref)
Yes	**1.64 (1.10-2.44)**

Hazard ratios and 95% confidence intervals in bold are statistically significant at p-value < 0.05

4.4. Sensitivity analysis

4.4.1. Components of CRC and ACOS

The Cox proportional hazard regression model was also run for each component of CRC (COPD and asthma) and for ACOS, with the results presented in **Table 4.4**. For the purpose of comparison, I used the same variables as in the CRC model. No SES indicator was statistically significant when asthma or ACOS were used as the dependent variable. The number of new cases that developed at follow-up in the case of asthma and ACOS may be insufficient to generate enough power to detect significant effects. The COPD model result, however, was similar to the CRC model, but with a slightly higher HR. For instance, individuals in lower and lower-middle income adequate households had HRs of 2.62 (95% CI: 1.23 – 5.58) and 1.95 (95% CI: 1.22 – 3.12), respectively for the COPD model, whereas those for the CRC model were 2.22 (95% CI: 1.01 – 4.89) and 1.66 (95% CI: 1.08 – 2.56), respectively. In addition, the HR for current smokers was statistically significant in the COPD model but not in the CRC model.

Table 4. 4. Multivariable Cox proportional hazard model results for COPD, asthma and ACOS

Variables	COPD HR (95% CI)	Asthma HR (95% CI)	ACOS HR (95% CI)
Age group			
(18-64) years	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
65+ years	1.34 (0.94-1.91)	1.11 (0.62-1.99)	1.67 (0.83-3.34)
Sex			
Male	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Female	1.13 (0.81-1.57)	1.40 (0.82-2.39)	1.01 (0.52-1.97)

Variables	COPD HR (95% CI)	Asthma HR (95% CI)	ACOS HR (95% CI)
Quadrants			
Southwest	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Southeast	1.54 (0.90-2.65)	1.97 (0.89-4.38)	2.87 (0.93-8.88)
Northeast	1.07 (0.62-1.85)	1.32 (0.57-3.04)	1.87 (0.58-6.08)
Northwest	1.58 (0.90-2.75)	0.99 (0.41-2.38)	1.99 (0.53-7.52)
MIZ			
No MIZ	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Weak MIZ	1.00 (0.69-1.43)	1.46 (0.76-2.82)	**0.52 (0.28-0.97)**
Moderate MIZ	**0.48 (0.25-0.91)**	2.09 (0.81-5.40)	0.34 (0.09-1.38)
Education			
Less than high school	1.33 (0.86-2.05)	1.56 (0.82-2.96)	1.55 (0.65-3.67)
High school	1.09 (0.76-1.55)	1.19 (0.65-2.18)	1.54 (0.78-3.06)
Postsecondary	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Financial strain			
Not enough money	0.87 (0.55-1.38)	1.14 (0.59-2.23)	0.76 (0.33-1.71)
Just enough money	1.16 (0.80-1.71)	0.72 (0.36-1.45)	0.96 (0.45-2.08)
Some money	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Household income adequacy			
Lower income	**2.62 (1.23-5.58)**	1.51 (0.34-6.76)	3.11 (0.93-10.43)
Lower middle income	**1.95 (1.22-3.12)**	1.55 (0.77-3.11)	2.13 (0.90-5.05)
Upper middle income	1.20 (0.83-1.73)	1.09 (0.61-1.95)	1.17 (0.55-2.47)
High income	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Molds			
Yes	0.84 (0.60-1.19)	0.83 (0.48-1.46)	0.76 (0.38-1.55)
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Natural gas			
Yes	**1.42 (1.02-1.97)**	1.09 (0.64-1.86)	0.60 (0.29-1.25)
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Crowding index			
Less than 1	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Between 1 and 2	0.78 (0.55-1.11)	1.63 (0.97-2.72)	1.09 (0.56-2.10)
Greater than 2	1.40 (0.15-12.71)	0.00 (0.00-0.00)	0.00 (0.00-0.00)
Home cigarette smoke			
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	1.15 (0.68-1.95)	0.68 (0.27-1.74)	1.62 (0.70-3.75)
Wood dust			
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	1.09 (0.76-1.57)	0.95 (0.55-1.63)	1.11 (0.58-2.11)
Asbestos dust			
No	0.70 (0.41-1.18)	1.27 (0.48-3.33)	0.52 (0.23-1.18)
Yes	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Mine dust			
No	0.75 (0.43-1.33)	0.48 (0.20-1.15)	1.05 (0.33-3.41)

Variables	COPD HR (95% CI)	Asthma HR (95% CI)	ACOS HR (95% CI)
Yes	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Smoking status			
Current smoker	**2.00 (1.08-3.70)**	0.73 (0.21-2.59)	0.96 (0.32-2.88)
Ex-smoker	**1.83 (1.29-2.59)**	1.13 (0.67-1.90)	0.92 (0.46-1.86)
Non-smoker	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
BMI			
Normal (< 25)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Overweight (>=25 and <=30)	1.41 (0.97-2.07)	1.13 (0.64-1.98)	1.07 (0.48-2.37)
Obese (>30)	1.05 (0.69-1.59)	1.13 (0.59-2.14)	1.54 (0.67-3.55)
1st year of life on farm			
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	1.21 (0.85-1.73)	0.69 (0.43-1.12)	0.94 (0.50-1.79)
At least 1 parent had lung disease			
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	**1.69 (1.24-2.32)**	1.26 (0.75-2.11)	1.75 (0.92-3.32)
Any allergy			
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	**1.51 (1.08-2.10)**	**3.33 (1.95-5.67)**	**4.83 (2.40-9.70)**
Heart disease			
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	1.35 (0.80-2.29)	0.41 (0.09-1.87)	0.87 (0.28-2.65)
Diabetes			
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	**1.76 (1.18-2.63)**	0.94 (0.36-2.49)	1.28 (0.51-3.25)

Hazard ratios and 95% confidence intervals in bold are statistically significant at p-value < 0.05

CHAPTER 5
DISCUSSION

This study used data from the Saskatchewan Rural Health Study (SRHS) to determine the incidence of CRC and in relation to SES. The chapter begins with a summary of the main results, followed by an integration of findings with the relevant research literature. A discussion of the study's strengths and limitations is then provided, followed by study implications, recommendations, and conclusions.

5.1. Main findings

The overall cumulative incidence of CRC was 7.1% between 2010 and 2014, and among those aged 65 years and older, 9.1%. No sex differences in CRC incidence emerged. The 4-year incidence of COPD, asthma, and ACOS was 6.2%, 2.2%, and 1.5%, respectively. In multivariable analysis, lower household income adequacy was associated with an increased risk of CRC.

5.1.1. Incidence of CRC

Studies estimating the cumulative incidence of CRC as defined in this study are scarce; therefore, the incidence of components of CRC, COPD and asthma, were compared to existing studies. According to the CCDSS report [7], the incidence of asthma (age 15+ yrs) and COPD (age 35+ yrs) in 2011/2012 was approximately 0.4% and 1.2%, respectively,[2] compared to the annual incidences of 0.6% (18yrs+) and 1.5% (35yrs+) observed in this study. Although these comparisons are based on somewhat imprecise estimates[2], they do suggest the possibility that the incidence of asthma, and especially COPD, may be somewhat higher in rural Saskatchewan than in the general Canadian population. A higher proportion of older people and/or a higher prevalence of some CRC risk factors in rural than general Canadian populations might contribute to these disparate estimates[32,33]. However, it is important to note that the CCDSS estimates are based on administrative health services data, and those of the present study, on self-report survey data. Also using administrative data, but restricted to the Albertan general

[2] These were estimated by averaging the age-specific incidence rates for asthma (Fig 1.7, p.13) and COPD (Fig 2.5, p.30) provided in the CCDSS report. The corresponding 2011/12 annual age-standardized rates in Canada were 0.5% for asthma and 0.9% for COPD.

population, a recent study[173] estimated the crude incidence of asthma in 2015 for those age 15 years and older at 0.5%.[3] Due to methodological variations, any comparisons of incidence between studies need to be done with caution.

5.1.2. SES and incidence of CRC

Lower SES in the present study was associated with an increased risk of CRC; that is, individuals with lower household income in 2010 were more likely than their higher income counterparts to report the development of CRC four years later. Lower educational level, but not financial strain, was associated with increased CRC risk in unadjusted analysis; however, statistical significance was not attained in the multivariable phase for both variables.

Numerous reviews have been published in recent years that conclude lower SES is associated with an elevated occurrence of CRC [19,125,132], which is in support of the present findings. However, for the most part, such claims have been made on the basis of cross-sectional studies examining prevalent CRC. The use of prevalence results in a lack of clarity concerning whether SES is associated with disease development, duration, or some combination of the two [174]. Very few studies have examined the relationship between SES and incidence of CRC; of the four that could be located in the last 10 years [51,150–152], one of them analyzed a subset of the data used in this study (incident bronchitis) [150], and reported similar results. A Danish birth cohort study of individuals born between 1961 and 1971 and followed from age 30 years to the onset of COPD reported an inverse relationship between indicators of early life SES (parental income, education and occupation) and incidence of COPD, but with parent education identified as exerting the strongest influence on COPD risk [152]. Although early life SES was the focus of this study, a considerable amount of COPD incidence was mediated via the social position participants attained in adulthood. In a recent Korean cohort study of 40-69 year olds, Kim and colleagues found that both lower income and education increased COPD risk, but only education remained statistically significant in adjusted analysis.[51] In a prospective cohort study of black American women, lower individual-level education, but not area-level SES, was associated with an elevated risk of adult asthma in adjusted analysis; individual income was not examined in the study[151]. Finally, one earlier Canadian study used 1994/5-1996/7 longitudinal data from the

[3] This estimate was calculated based on supplementary raw data for those 15 years of age and older (Table S3, 2015) provided by Bosonea et al.[173]

National Population Health Survey (NPHS) and found that income adequacy was not associated with the incidence of asthma, which the authors suggest may be due to limited statistical power.[175]

The longitudinal studies reviewed above are generally in line with the results of this study; that is, that lower SES is associated with an increased risk of CRC. What does vary, however, is the particular SES indicator of most importance. In this study, household income adequacy was most relevant, whereas, in the studies reviewed above, educational attainment appeared most germane. The choice of SES indicator is not simply an academic question, as these measures are believed to lead to poorer health through related but different pathways [176], which in turn, have implications for the development of appropriate interventions.

However, many challenges exist when trying to compare results across studies due to differences in country of origin, the demographic characteristics of participants, the measures used, and the statistics applied, including the variables used in adjusted analyses. One major difference between all these studies and the current study is the population of focus – they are not rural studies. Rurality may create subtle differences in the results due to the uniqueness of rural demographics [39]. The insignificant association of education with CRC among rural dwellers in this study may be explained by the role of education as an SES indicator and the structure of rural employment and income. Education is a strong determinant of future employment and income and captures the transition from a parent's SES to the individual's own SES in adulthood [177,178]. Compared to urban centres, rural areas are usually characterised by more self-employment and intergenerational succession, and in this context, knowledge may be acquired through watching and doing [179]. Therefore, education may be a less relevant determinant of future employment and income among some rural dwellers. It is important to note that, in contrast to the present study, an earlier study using baseline SRHS data[27] reported an association between lower education and an increased odd of CRC (consisting of COPD and bronchitis), especially among women less than 65 years of age. The discrepant results, however, may be due to the cross-sectional design of that earlier study, as well as differences in the measurement of CRC and control of confounders.

5.1.2.1 How lower income may lead to higher CRC incidence

The effect of income on health is well documented in the literature, with lower-income individuals having poorer health [180–182]. The household income patterning of CRC can be explained by the fact that those with lower income may be more likely to be exposed to established risk factors for CRC, a mechanism referred to as differential exposure [37]. Research has shown that poverty can increase the exposure of people to behavioural risk factors, such as smoking, obesity, and substance use, which in turn may increase the risk of developing chronic health conditions such as COPD and asthma [134,137,138,183]. The prevalence of smoking, one of the main risk factors of CRC [135,136], continues to rise among socially disadvantaged groups [135,136]. There is also evidence of a relationship between lower income and subsequent obesity [122], possibly due to poorer nutrition and lack of physical activity. Risk behaviours may also be related to a negative social environment [143–145]. This social environment encompasses the communities in which we live, the organisations of our work and the policies created to order our lives [129]. A negative social environment may also be characterised by a higher prevalence of stressors, such as marital discord and high demands at work without adequate resources to meet such demands [184]. These stress-related factors are known to depress the immune function [185] and increase susceptibility to or exacerbate several diseases and disorders, including asthma [186] and other chronic respiratory diseases [187]. So even within a rural setting, those with lower income may do the most demanding work and yet receive the least pay and/or may also have marital issues that could stem from inadequate household income.

People with higher incomes can enjoy place-based health benefits. Thus, the conditions and assets in these people's environment positively influence their health [188]. Low-income individuals cannot afford such luxury and may be forced to settle for what they can afford. Low-income individuals live in unhealthy environments without good shelter, clean water, adequate sanitation [189–191], and a cleaner energy source for house heating, not by choice but by necessity. For instance, in this study, the proportion of individuals in higher income adequate households who used natural gas as a source of household heat is greater than those in lower income adequate households. Another mechanism through which low income dictates poor health is through the proportion of expenditure on health. Health expenditure often comprises a large proportion of lower income individuals' household income, increasing their economic burden and hindering their accessibility to life needs such as good nutrition and quality education [192,193].

In Canada, this health expenditure may not be direct due to the free health care provision. However, in the case of rural areas, the transportation to healthcare centres and the purchase of prescription drugs, which are not mostly covered by provincial insurance, may contribute to a significant proportion of income for lower income households. Good nutrition may resolve the issue of obesity, which is one of the risk factors of CRC [194]. However, a 2006 Canadian rural health reported indicated that rural people were lower consumers of fruit, and a greater proportion were smokers [39]. Health expenditure has become a significant cause of household poverty [195], and poor quality and ineffective healthcare, which characterised rurality, only exacerbate the burdens of lower-income groups [196]. They are also less likely to have extra health insurance or receive new drugs and technologies [133,134]. With these challenges, lower-income individuals may not have the urge to see a doctor [197]. Therefore, they may be less likely to be diagnosed and treated for early signs of chronic diseases.

Although these are plausible mechanisms, it is important to note that after adjusting for many risk factors in the multivariable analysis, lower income adequacy remained associated with an increase in CRC risk. This could be due to a number of reasons, including residual confounding. In addition, it is possible that exposures on the causal pathway between income and CRC that are of most relevance to rural populations were not measured in this study. Future research is needed to systematically assess and identify the material and psychosocial mechanisms linking lower income with the development of CRC within rural settings.

5.2. Other findings

Other statistically significant risk factors of this study included not using natural gas as a source of heat, being an ex-smoker, having at least one parent with lung disease, allergy, and diabetes.

This study found using natural gas as a heat source to be a protective factor for CRC incidence. A previous study using SRHS data also found natural gas as the source of heat to be a protective factor against chronic bronchitis but only in unadjusted analyses [150]. With the use of natural gas as the heat source, there is less possibility of exposure to smoke, which could explain the protective effect detected. Contrary to this finding, Rabinowtz et al. found that more people from households in close proximity with natural gas wells reported respiratory symptoms than those distanced from the well [198].

Smoking is generally recognised as a major risk factor of CRC [199–204]. In this study, I found that ex-smoking was associated with a higher risk of CRC incidence, but not current smoking. Previous research has found both current and ex-smoking to be associated with the incidence of asthma [199–201] and COPD [202–204]. For example, in the European Community Respiratory Health Survey [201] the incidence rate ratios for current and ex-smoking compared with never smoking were 1.3 (95% CI, 1.0 – 1.6) and 1.4 (95% CI, 1.0 – 2.0), respectively. A 10-year longitudinal study in Northern Sweden using the third survey of the Obstructive Lung Disease in Northern Sweden studies cohort also found current smoking and ex-smoking to be associated with COPD incidence. The lack of significant association between current smokers and CRC in this study may be due to the relatively short follow-up time. Also, ex-smokers may be individuals who quit because of worsening health conditions. The short follow-up period may explain why this study did not find current smoking to be statistically significant in relation to incidence of CRC.

Individuals with a family history of CRC had a 1.54 (95% CI: 1.14 – 2.04) times higher risk of CRC than those without such a history. This finding is consistent with a rural Chinese study that also found individuals with a family history of COPD to have 1.61 times higher odds of COPD after a 4-year follow-up [194]. Family history as a risk factor for CRC may be explained by genetic and environmental mechanisms. Genetically, parents with lung disease-carrying genes may pass it on to their children, increasing the risk of CRC [12,100]. Environmentally, parents who smoke may expose their children to second-hand smoke, and/or children of smoking parents may be more likely to also smoke during their lifetime compared to children of non-smoking parents. [205].

Individuals with any allergy in the study were also at higher risk (HR=1.66; 95% CI: 1.22 – 2.27) of CRC than those without it. Bui and colleges found that the trajectories of allergies were associated with varying risk of COPD [206]. Allergies may lead to CRC through weakening of the immune system and the increase in mucus production [207]. This may lead to exacerbation and decline in in lung function [208]. This study also found diabetes to be a risk factor for CRC. The result of this study is consistent with previous studies [209,210]. After adjusting for age, sex, race/ethnicity, smoking status, BMI, educational attainment, alcohol consumption, and outpatient visits, Ehrlich et al. found that individuals with diabetes have an increased risk of asthma, COPD, fibrosis, and pneumonia [209]. Thomsen and colleagues observed an almost double asthma

risk in patients with type II diabetes [210]. This may be explained by the higher prevalence of obesity, a risk factor for both diabetes and CRC [211,212], among people with diabetes. Another possible explanation may be the effect of diabetes on the severity and clinical course of CRC [213]. High blood sugar level can affect the blood vessels in the lungs leading to lung issues or worsen lung conditions [214].

5.3. Sensitivity analysis results

After analysing the disaggregated CRC components, the results indicate that COPD highly influenced the estimates of this study, as the COPD model yielded similar estimates. .In other words, asthma cases may have contributed less to the CRC model. These findings were not entirely surprising given the short follow-up period and the fact that asthma is more prevalent in children than adults [8]. The four years follow-up period may be too short to permit the onset of new adult asthma cases. The issue of new asthma cases also translates to the ACOS model, where the number of new cases was even less, given the definition of ACOS in this study.

5.4. Study strengths and limitations

To the best of my knowledge, this study is one of the few studies to examine the longitudinal association of SES and CRC in a rural population in North America. The longitudinal nature of the study gives a better understanding of the SES patterning of CRC among rural dwellers in Saskatchewan. Previous cross-sectional studies have demonstrated a relationship between lower SES components of CRC [27,215], but poor health may also result in lower SES, making it difficult to make valid conclusions based on findings from cross-sectional studies. Longitudinal studies are able to address the issue of reverse causality between SES and disease and establish temporality, which fulfils one of the criteria for causality. Therefore, this longitudinal study provides evidence that lower income leads to a higher incidence of CRC among rural dwellers. The study estimates are generally consistent with previous studies conducted at different places, times, and in different populations [19,125,132]. The highest HR estimate for SES is 2.22, which can be considered a moderately strong association between SES and CRC; however, magnitude alone may not be an appropriate measure of strength of association and thus, evidence of causality [216]. SES, especially income adequacy, also demonstrated a dose-response relationship with CRC; that is, as individuals descended the SES

spectrum, the risk of CRC increased. Although only income adequacy was associated with CRC in multivariable analysis, in univariate analysis, a dose-response relationship was also evident with education and financial strain. Finally, the results observed in the present study are consistent with social determinants of health frameworks[37] which detail plausible mechanisms through which lower SES may lead to a higher risk of CRC.

The results of the present study also point to the importance of the social determinants of respiratory health in a rural population. In North America, the rural health literature is populated with service accessibility. [39] With only a few exceptions[27,34–36,217,218], when SES factors are included in studies, they are rarely used as primary exposures of interest. Considering that there may be substantial variation in the material and psychosocial resources that underlie SES-health gradients between urban and rural areas [39], the SES-health patterns observed in general populations might not necessarily be accurate in rural settings[34–36]. Therefore, this study draws attention to the importance of socioeconomic circumstances as determinants of respiratory health in rural residents. It highlights the fact that the issue of rural health should not only be centred on quality and accessibility of services, but also, on the social determinants of respiratory health inequities.

Study weaknesses must also be acknowledged. The first has to do with the definition of CRC. CRC was defined as having one or both COPD and asthma. These two diseases are clinically and epidemiologically different [219]. This is evident in this study's sensitivity analysis, whereby household income adequacy was a significant risk factor in the COPD model but not in the asthma model. The definition used in this study, however, has been used in a previous study to estimate the association between SES and chronic lung diseases [220]. A related limitation is how each component of CRC was measured in the present study; that is, each was self-reported, whereas other studies have used physical examinations [7,147] or administrative data [173]. However, self-reported data have been shown to generate valid and reliable measures of chronic diseases [221,222].

Second, out of the eligible participants, only 42% responded in the baseline survey [153,154]. If the respondents and non-respondents were systematically different, incidence and HR estimates may be under- or over-estimated. Previous research has shown that non-respondents are more likely to have poor health and belong to lower SES than respondents [223]. This is likely to result in underestimation of the HR. Also, out of those who participated in the baseline survey,

only 62.65% participated in the follow-up survey [153,154], introducing a likely self-selection bias where lower income individuals and those who may have developed the disease are more likely to drop out of the study than higher income individuals. This may have resulted in an underestimation of the incidence of CRC and the HR for lower income.

Even though several potential confounders have been adjusted for, the existence of residual confounding cannot be ruled out. Given that chronic respiratory conditions, especially COPD, develop slowly over the years, where an individual lived in the first year of life and whether parents smoked during pregnancy may not provide sufficient adjustment for early life exposures. On the other hand, some variables, such as smoking status, may lie in the SES-CRC causal pathway. The inclusion of these variables may result in an underestimation of the HR for the SES indicators.

Also, the study lacked ethnic diversity due to very few Indigenous participants, which hindered the exploration of how ethnicity might intersect with SES in a rural setting. The study population may also not be a representative sample of rural Saskatchewan. The sample for this study covered a larger proportion of older people than the population covered by Saskatchewan Health, including Saskatchewan residents registered for provincial health coverage [224]. Therefore, caution must be taken when generalising results related to CRC and its components to the entire Saskatchewan rural population.

5.5. Recommendations and future research

Rurality itself may contribute to or aggravate SES disparities in respiratory health. Therefore, understanding how the SES of rural dwellers affects their health is vital if policies and interventions are to engineer a significant improvement in the well-being of rural dwellers. Usually, rural health policies are predicated on the assumption that rural health is worse than urban health [225,226], and that rurality may be responsible for such worse health outcomes. However, this study demonstrated that the SES of rural people is very important in determining their health outcomes. Many rural policies have focussed on improving health care services through innovative models of service delivery for rural communities, including telehealth and initiatives to retain healthcare providers in rural areas [227–229]. However, focusing solely on distance and access without addressing the fundamental social and structural processes operating in the broader context may lead to sub-optimal policy interventions [229].

To improve on these policies, the issue of poverty, inequality, inequities in resource allocation and discrimination must be addressed equally in rural communities as attended to in urban settings. The greatest health gains may be realised among lower SES groups, but designed interventions are, unfortunately, not directed towards these groups [230]. Rural health policies must consider the difference in people's social positions even within the rural community without assuming that solving the problem of distance and access resolves the problem for all rural dwellers equally.

Given these study limitations, additional research is needed with a longer follow-up period and a larger number of participants to allow for CRC disaggregation in relation to SES. As explained earlier, COPD and asthma are clinically different and may yield different results with a larger sample. As observed in this study, SES was not associated with asthma in the sensitivity analysis, but this does not mean that SES isn't a determinant of asthma in rural people. COPD was also observed to drive the estimates from the CRC model. This highlights the difference in estimates that may arise when larger numbers are available.

Secondly, there are economic and demographic differences between provinces and territories in Canada; therefore, a broader study covering most rural areas in Canada would facilitate result generalisation. Also, a more ethnically diverse rural study population would allow for a more nuanced examination of SES inequities in the incidence of CRC.

5.6. Conclusions

This study used the Cox proportional hazard model to estimate the association between SES and CRC in rural Saskatchewan. The incidence of CRC during the 4 years follow-up period was 7.01% and increased with decreasing SES. Household income adequacy was inversely associated with CRC incidence in multivariable analyses, whereas education and financial strain were not related. The differential exposure mechanism could explain the income-CRC pathway observed, whereby people of lower SES are more likely to be exposed to risk factors for CRC, such as smoking, poor household environments, and harmful workplace exposures. Despite education being an important SES indicator in the literature [151,152], it may not be as relevant among rural Canadians in this study. Also, rural communities may have strong community bonds so that individuals can seek emergency help from each other [39], thereby limiting any damaging impact of financial strain.

In conclusion, rural dwellers should not be treated as a unit when addressing their health challenges, as these challenges are worse for some groups than others based on socioeconomic status.

REFERENCES

1. World Health Organisation. Chronic Respiratory Diseases. Available at: https://www.who.int/health-topics/chronic-respiratory-diseases#tab=tab_1. (Accessed: 5th May 2021)
2. The Lancet. GBD 2017: A Fragile World. *Lancet* **392**, (2018).
3. Public Health Agency of Canada. Chronic Respiratory Diseases . Available at: https://www.canada.ca/en/public-health/services/chronic-diseases/chronic-respiratory-diseases.html. (Accessed: 5th May 2021)
4. Soriano, J. B. *et al.* Prevalence and Attributable Health Burden of Chronic Respiratory Diseases, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet Respir. Med.* **8**, 585–596 (2020).
5. World Health Organization. The Top 10 Causes of Death. (2020). Available at: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. (Accessed: 17th February 2021)
6. Statistics Canada. Table 13-10-0096-19 Chronic Obstructive Pulmonary Disease (COPD), 35 Years and Over. doi:https://doi.org/10.25318/1310009601-eng
7. Public Health Agency of Canada. *Report from the Canadian Chronic Disease Surveillance System: Asthma and Chronic Obstructive Pulmonary Disease (COPD) in Canada, 2018*. (2018).
8. Pearce, N. *et al.* Worldwide Trends in the Prevalence of Asthma Symptoms: Phase III of the International Study of Asthma and Allergies in Childhood (ISAAC). *Thorax* **62**, 757–765 (2007).
9. To, T. *et al.* Global Asthma Prevalence in Adults: Findings from the Cross-sectional World Health Survey. *BMC Public Health* **12**, 1–8 (2012).
10. European Respiratory Society. The Economic Burden of Lung Disease . in *European Lung White Book*
11. Patra, J. *et al. Economic Cost of Chronic Disease in Canada: 1995-2003*. (2007).
12. European Respiratory Society. Genetic Susceptibility. in *European Lung White Book*
13. Chen, W., Thomas, J., Sadatsafavi, M. & FitzGerald, J. M. Risk of Cardiovascular Comorbidity in Patients with Chronic Obstructive Pulmonary Disease: A Systematic Review and Meta-Analysis. *Lancet Respir. Med.* **3**, 631–639 (2015).
14. Ospina, M. B. *et al.* Incidence and Prevalence of Chronic Obstructive Pulmonary Disease among Aboriginal Peoples in Alberta, Canada. *PLoS One* **10**, e0123204 (2015).
15. Postma, D. S., Bush, A. & Van Den Berge, M. Risk factors and early Origins of Chronic Obstructive Pulmonary Disease. *Lancet* **385**, 899–909 (2015).
16. Van Remoortel, H. *et al.* Risk Factors and Comorbidities in the Preclinical Stages of

Chronic Obstructive Pulmonary Disease. *Am. J. Respir. Crit. Care Med.* **189**, 30–38 (2014).

17. Vozoris, N. T. & O'Donnell, D. E. Prevalence, Risk Factors, Activity Limitation and Health Care Utilization of an Obese Population-Based Sample with Chronic Obstructive Pulmonary Disease. *Can. Respir. J.* **19**, e18–e24 (2012).

18. Eisner, M. D. *et al.* Socioeconomic Status, Race and COPD Health Outcomes. *J. Epidemiol. Community Health* **65**, 26–34 (2011).

19. Gershon, A. S., Dolmage, T. E., Stephenson, A. & Jackson, B. Chronic obstructive Pulmonary Disease and Socioeconomic Status: A systematic Review. *COPD J. Chronic Obstr. Pulm. Dis.* **9**, 216–226 (2012).

20. Grigsby, M. *et al.* Socioeconomic Status and COPD Among Low-and Middle-Income Countries. *Int. J. Chronic Obstr. Pulm. Dis.* **11**, 2497–2507 (2016).

21. Kanervisto, M. *et al.* Low Socioeconomic Status is Associated with Chronic Obstructive Airway Diseases. *Respir. Med.* **105**, 1140–1146 (2011).

22. Lange, P., Marott, J. L., Vestbo, J., Ingebrigtsen, T. S. & Nordestgaard, B. G. Socioeconomic Status and Prognosis of COPD in Denmark. *COPD J. Chronic Obstr. Pulm. Dis.* **11**, 431–437 (2014).

23. O'neil, A., Russell, J. D., Thompson, K., Martinson, M. L. & Peters, S. A. E. The Impact of Socioeconomic Position (SEP) On Women's Health Over the Lifetime. *Maturitas* **140**, 1–7 (2020).

24. *Social Determinants of Health: The Solid Facts*. (World Health Organization, 2003).

25. Commission on Social Determinants of Health. *Closing the Gap in a Generation: Health Equity Through Action on the Social Determinants of Health*. (2008).

26. Zhang, D.-D. *et al.* Association between Socioeconomic Status and Chronic Obstructive Pulmonary Disease in Jiangsu Province, China: A Population-based Study. *Chin. Med. J. (Engl).* **134**, 1552–1560 (2021).

27. Janzen, B. *et al.* Exploring Diversity in Socioeconomic Inequalities in Health among Rural Dwelling Canadians. *J. Rural Heal.* **31**, 186–198 (2015).

28. Li, X. *et al.* Association of Socioeconomic and Lifestyle Factors with Chronic Non-communicable Diseases and Multimorbidity among the Elderly in Rural Southwest China. *J. Public Health (Oxf).* **42**, 239–246 (2020).

29. Kainu, A. *et al.* COPD in Helsinki, Finland: Socioeconomic Status Based on Occupation has an Important Impact on Prevalence. *Scand. J. Public Health* **41**, 570–578 (2013).

30. Schyllert, C. *et al.* Low socioeconomic status relates to asthma and wheeze, especially in women. *ERJ Open Res.* **6**, 00258–02019 (2020).

31. Health Canada. *Respiratory Disease in Canada*. (2001).

32. Karunanayake, C. P., Rennie, D. C., Pahwa, P., Chen, Y. & Dosman, J. A. Predictors of

Respiratory Symptoms in a Rural Canadian Population: A Longitudinal Study of Respiratory Health. *Can. Respir. J.* **18**, 149–153 (2011).

33. Matthews, K. A. *et al.* Health-Related Behaviors by Urban-Rural County Classification — United States, 2013. *MMWR. Surveill. Summ.* **66**, 1–8 (2017).

34. Hayward, M. D., Pienta, A. M. & McLaughlin, D. K. Inequality in Men's Mortality: The Socioeconomic Status Gradient and Geographic Context. *J. Health Soc. Behav.* **38**, 313–330 (1997).

35. Pickett, W. *et al.* Socioeconomic Status and Injury in a Cohort of Saskatchewan Farmers. *J. Rural Heal.* **27**, 245–254 (2011).

36. Wanless, D., Mitchell, B. A. & Wister, A. V. Social Determinants of Health for Older Women in Canada: Does Rural–Urban Residency Matter? *Can. J. Aging* **29**, 233–247 (2010).

37. Solar, O. & Irwin, A. *A Conceptual Framework for Action on the Social Determinants of Health. Social Determinants of Health Discussion Paper 2 (Policy and Practice)*. *World Health Organization Geneva* (2010).

38. Raju, S. *et al.* Rural Residence and Poverty are Independent Risk Factors for Chronic Obstructive Pulmonary Disease in the United States. *Am. J. Respir. Crit. Care Med.* **199**, 961–969 (2019).

39. Canadian Institute for Health Information. *How Healthy Are Rural Canadians?* (2006).

40. Public Health Ontario. Chronic Respiratory Diseases . (2021). Available at: https://www.publichealthontario.ca/en/diseases-and-conditions/chronic-diseases-and-conditions/chronic-respiratory-diseases. (Accessed: 30th June 2021)

41. Eisner, M. D., Yelin, E. H., Trupin, L. & Blanc, P. D. The Influence of Chronic Respiratory Conditions on Health Status and Work Disability. *Am. J. Public Health* **92**, 1506–1513 (2002).

42. Society, A. T. Standards for the Diagnosis and Care of Patients with Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* **152**, 77S-120S (1995).

43. Siafakas, N. *et al.* Optimal Assessment and Management of Chronic Obstructive Pulmonary Disease (COPD). The European Respiratory Society Task Force. *Eur. Respir. J.* **8**, 1398–1420 (1995).

44. Vestbo, J. *et al.* Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease - GOLD Executive Summary. *Am. J. Respir. Crit. Care Med.* **187**, 347–365 (2013).

45. Swanney, M. P. *et al.* Using the Lower Limit of Normal for the FEV 1 /FVC Ratio Reduces the Misclassification of Airway Obstruction. *Thorax* **63**, 1046–1051 (2008).

46. Herrera, A. C. *et al.* COPD Underdiagnosis and Misdiagnosis in a High-Risk Primary Care Population in Four Latin American Countries. A Key to Enhance Disease Diagnosis: The PUMA Study. *PLoS One* **11**, e0152266 (2016).

47. Hill, K. *et al.* Prevalence and Underdiagnosis of Chronic Obstructive Pulmonary Disease among Patients at Risk in Primary Care. *Can. Med. Assoc. J.* **182**, 678 (2010).

48. Diab, N. *et al.* Underdiagnosis and Overdiagnosis of Chronic Obstructive Pulmonary Disease. *Am. J. Respir. Crit. Care Med.* **198**, 1130–1139 (2018).

49. Han, M. L. K. *et al.* Spirometry Utilization for COPD: How Do We Measure Up? *Chest* **132**, 403–409 (2007).

50. Maleki-Yazdi, M. R. *et al.* The Burden of Illness in Patients with Moderate to Severe Chronic Obstructive Pulmonary Disease in Canada. *Can. Respir. J.* **19**, 319–324 (2012).

51. Kim, C. Y. *et al.* Longitudinal Evaluation of the Relationship Between Low Socioeconomic Status and Incidence of Chronic Obstructive Pulmonary Disease: Korean Genome and Epidemiology Study (KoGES). *Int. J. Chron. Obstruct. Pulmon. Dis.* **15**, 3447–3454 (2020).

52. Evans, J., Chen, Y., Camp, P. G., Bowie, D. M. & Mcrae, L. Estimating the Prevalence of COPD in Canada: Reported Diagnosis Versus Measured Airflow Obstruction. *Heal. Rep.* **25**, 3–11 (2014).

53. Hamm, N. C. *et al.* Trends in Chronic Disease Incidence Rates from the Canadian Chronic Disease Surveillance System. *Heal. Promot. Chronic Dis. Prev. Canada* **39**, 216–224 (2019).

54. Leung, C. *et al.* The Prevalence of Chronic Obstructive Pulmonary Disease (COPD) and the Heterogeneity of Risk Factors in the Canadian Population: Results from the Canadian Obstructive Lung Disease (COLD) Study. *Int. J. Chron. Obstruct. Pulmon. Dis.* (2021). doi:10.2147/COPD.S285338

55. World Health Organization. Asthma. (2021). Available at: https://www.who.int/news-room/fact-sheets/detail/asthma. (Accessed: 11th October 2021)

56. Pavord, I. D. *et al.* After Asthma: Redefining Airways Diseases. *Lancet* **391**, 350–400 (2018).

57. Kleinert, S. & Horton, R. After Asthma: Airways Diseases Need a New Name and a Revolution. *Lancet* **391**, 292–294 (2018).

58. National Asthma Education and Prevention Program. *Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma*. (2007).

59. Busse, W. W. Asthma Diagnosis and Treatment: Filling in the Information Gaps. *J. Allergy Clin. Immunol.* **128**, 740–750 (2011).

60. Hederos, C. A., Hasselgren, M., Hedlin, G. & Bornehag, C. G. Comparison of Clinically Diagnosed Asthma with Parental Assessment of Children's Asthma in a Questionnaire. *Pediatr. Allergy Immunol.* **18**, 135–141 (2007).

61. Prescott, E., Lange, P., Vestbo, J. & The Copenhagen City Heart Study Group. Effect of Gender on Hospital Admissions for Asthma and Prevalence of Self-Reported Asthma: A Prospective Study based on a Sample of the General Population . *Thorax* **52**, 287–289

(1997).

62. Van Wonderen, K. . *et al.* Different Definitions in Childhood Asthma: How Dependable is the Dependent Variable? *Eur. Respir. J.* **36**, 48–56 (2010).

63. Jenkins, M. A. *et al.* Validation of Questionnaire and Bronchial Hyperresponsiveness against Respiratory Physician Assessment in the Diagnosis of Asthma. *Int. J. Epidemiol.* **25**, 609–616 (1996).

64. Pekkanen, J., Sunyer, J., Anto, J. M. & Burney, P. Operational Definitions of Asthma in Studies on its Aetiology. *Eur. Respir. J.* **26**, 28–35 (2005).

65. Hargreave, F. E. & Nair, P. The Definition and Diagnosis of Asthma. *Clin. Exp. Allergy* **39**, 1652–1658 (2009).

66. Hansen, S. *et al.* A Comparison of Three Methods to Measure Asthma in Epidemiologic Studies: Results from the Danish National Birth Cohort. *PLoS One* **7**, 36328 (2012).

67. Postma, D. S., Reddel, H. K., ten Hacken, N. H. T. & van den Berge, M. Asthma and Chronic Obstructive Pulmonary Disease: Similarities and Differences. *Clin. Chest Med.* **35**, 143–156 (2014).

68. Kim, T. B. *et al.* The Reality of an Intermediate Type between Asthma and COPD in Practice. *Respir. Care* **57**, 1248–1253 (2012).

69. Bujarski, S., Parulekar, A. D., Sharafkhaneh, A. & Hanania, N. A. The Asthma COPD Overlap Syndrome (ACOS). *Curr. Allergy Asthma Rep.* **15**, 1–9 (2015).

70. Braman, S. S. The Chronic Obstructive Pulmonary Disease-Asthma Overlap Syndrome. *Allergy Asthma Proc.* **36**, 11–18 (2015).

71. Tho, N. Van, Park, H. Y. & Nakano, Y. Asthma–COPD Overlap Syndrome (ACOS): A Diagnostic Challenge. *Respirology* **21**, 410–418 (2016).

72. Chen, F. J., Huang, X. Y., Liu, Y. L., Lin, G. P. & Xie, C. M. Importance of Fractional Exhaled Nitric Oxide in the Differentiation of Asthma–COPD Overlap Syndrome, Asthma, and COPD. *Int. J. Chron. Obstruct. Pulmon. Dis.* **11**, 2390 (2016).

73. Bonten, T. N. *et al.* Defining Asthma-COPD Overlap Syndrome: A Population-based Study. *Eur. Respir. J.* **49**, 1–11 (2017).

74. Roth, G. A. *et al.* Global, Regional, and National Age-Sex-Specific Mortality for 282 Causes of Death in 195 Countries and Territories, 1980–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1736–1788 (2018).

75. World Health Organization. Global Surveillance, Prevention and Control of Chronic Respiratory Diseases: A Comprehensive Approach. (2012). Available at: https://www.who.int/gard/publications/GARD_Manual/en/. (Accessed: 5th May 2021)

76. James, S. L. *et al.* Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 354 Diseases and Injuries for 195 Countries and Territories, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789–1858 (2018).

77. Kyu, H. H. *et al.* Global, Regional, and National Disability-Adjusted Life-Years (Dalys) for 359 Diseases and Injuries and Healthy Life Expectancy (HALE) for 195 Countries and Territories, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1859–1922 (2018).

78. Ntritsos, G. *et al.* Gender-Specific Estimates of COPD Prevalence: A Systematic Review and Meta-analysis. *Int. J. Chron. Obstruct. Pulmon. Dis.* **13**, 1514 (2018).

79. Innes Asher, M. *et al.* Trends in worldwide asthma prevalence Number 1 in the series 'Innovations in asthma and its treatment' Edited by. *Eur. Respir. J.* **56**, (2020).

80. Moorman, J. E., Zahran, H., Truman, B. I., Molla, M. T. & Centers for Disease Control and Prevention (CDC). Current Asthma Prevalence - United States, 2006--2008. *MMWR Suiveillance Summ.* **Suppl**, 84–86 (2011).

81. Murray, C. J., Lopez, A. D., Mathers, C. D. & Stein, C. *The Global Burden of Disease 2000 Project: aims, Methods and Data Sources*. (2001).

82. Vos, T. *et al.* Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 328 Diseases and Injuries for 195 Countries, 1990–2016: A Systematic Analysis for the Global Burden of Disease Study 2016. *Lancet* **390**, 1211–1259 (2017).

83. Vos, T. *et al.* Years Lived with Disability (Ylds) for 1160 Sequelae Of 289 Diseases and Injuries 1990-2010: A Systematic Analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2163–2196 (2012).

84. Vos, T. *et al.* Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 301 Acute and Chronic Diseases and Injuries in 188 Countries, 1990-2013: A Systematic Analysis for the Global Burden of Disease Study 2013. *Lancet* **386**, 743–800 (2015).

85. Soriano, J. B. *et al.* Global, Regional, and National Deaths, Prevalence, Disability-Adjusted Life Years, and Years Lived with Disability for Chronic Obstructive Pulmonary Disease and Asthma, 1990–2015: A Systematic Analysis for the Global Burden of Disease Study 2015. *Lancet Respir. Med.* **5**, 691–706 (2017).

86. Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L. & Torrance, G. W. *Methods for the Economic Evaluation of Health Care Programmes* . *Oxford University Press* (Oxford University Press, 2015).

87. Ehteshami-Afshar, S., FitzGerald, J. M., Doyle-Waters, M. M. & Sadatsafavi, M. The Global Economic Burden of Asthma and Chronic Obstructive Pulmonary Disease. *Int. J. Tuberc. Lung Dis.* **20**, 11–23 (2016).

88. Viegi, G., Maio, S., Fasola, S. & Baldacci, S. Global Burden of Chronic Respiratory Diseases. *J. Aerosol Med. Pulm. Drug Deliv.* **33**, 171–177 (2020).

89. Osman, S., Ziegler, C., Gibson, R., Mahmood, R. & Moraros, J. The Association Between Risk Factors and Chronic Obstructive Pulmonary Disease in Canada: A Cross-sectional Study Using the 2014 Canadian Community Health Survey. *Int. J. Prev. Med.* **8**, (2017).

90. Kozyrskyj, A. L. & Becker, A. B. Rural-urban Differences in Asthma Prevalence:

Possible Explanations. *J. Allergy Clin. Immunol.* **113**, S306 (2004).

91. Manfreda, J. *et al.* Prevalence of Asthma Symptoms among Adults Aged 20-44 years in Canada. *Can. Med. Assoc. J. = J. L'association Medicale Can.* **164**, 995–1001 (2001).

92. Rennie, D. C. *et al.* Early Farm Residency and Prevalence of Asthma and Hay Fever in Adults. *J. Asthma* **53**, 2–10 (2016).

93. Tan, W. C. *et al.* Can Age and Sex Explain the Variation in COPD Rates Across Large Urban Cities? A Population Study in Canada. *Int. J. Tuberc. Lung Dis.* **15**, 1691–1696 (2011).

94. Public Health Agency of Canada. Economic Burden of Illness in Canada, 2010. (2017). Available at: https://www.canada.ca/en/public-health/services/publications/science-research-data/economic-burden-illness-canada-2010.html#res1. (Accessed: 10th May 2021)

95. Saskatchewan Ministry of Health. *Prevalence of Asthma, COPD, Diabetes, and Hypertension in Saskatchwewan, 2010/11*. (2013).

96. Saskatchewan Ministry of Health. *Incidence of asthma, COPD, diabetes, and hypertension in Saskatchewan, 2010/11*. (2013).

97. Government of Saskatchewan. *COPD in Saskatchewan (2014/15)*. (2015).

98. Saskatchewan Ministry of Health. Health Status Reports . Available at: https://www.saskatchewan.ca/government/government-structure/ministries/health/other-reports/health-status-reports. (Accessed: 13th May 2021)

99. Statistics Canada. Population Estimates on July 1st, by Age and Sex: Table: 17-10-0005-01. (2020). Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000501. (Accessed: 24th January 2021)

100. Alsamri, M. T. *et al.* Genetic Variants in Children with Chronic Respiratory Diseases. *Pediatr. Pulmonol.* **55**, 2389–2401 (2020).

101. Gabryelska, A., Kuna, P., Antczak, A., Białasiewicz, P. & Panek, M. IL-33 Mediated Inflammation in Chronic Respiratory Diseases—Understanding the Role of the Member of IL-1 Superfamily. *Front. Immunol.* **10**, (2019).

102. European Respiratory Society. Tobacco Smoking. in *European Lung White Book*

103. Ferkol, T. & Schraufnagel, D. The Global Burden of Respiratory Disease. *Ann Am Thorac Soc* **11**, 404–406 (2014).

104. GARD. *Global Surveillance, Prevention and Control of Chronic Respiratory Diseases: A comprehensive Approach*. (World Health Organization, 2007).

105. Van Gemert, F. *et al.* Prevalence of Chronic Obstructive Pulmonary Disease and Associated Risk Factors in Uganda (FRESH AIR Uganda): A Prospective Cross-sectional Observational Study. *Lancet Glob. Heal.* **3**, e44–e51 (2015).

106. Bousquet, J., Dahl, R. & Khaltaev, N. Global Alliance Against Chronic Respiratory Diseases. *Eur. Respir. J.* **29**, 233–239 (2007).

107. Bakke, P. S., Baste, V., Hanoa, R. & Gulsvik, A. Prevalence of Obstructive Lung Disease in a General Population: Relation to Occupational Title and Exposure to Some Airborne Agents. *Thorax* **46**, 863–870 (1991).

108. Tan, W. C. *et al.* Characteristics of COPD in Never-smokers and Ever-smokers in the General Population: Results from the CanCOLD Study. *Thorax* **70**, 822–829 (2015).

109. Salvi, S. S. & Barnes, P. J. Chronic Obstructive Pulmonary Disease in Non-smokers. *Lancet* **374**, 733–743 (2009).

110. Hagstad, S. *et al.* COPD among Non-smokers – Report from the Obstructive Lung Disease in Northern Sweden (OLIN) Studies. *Respir. Med.* **106**, 980–988 (2012).

111. Shin, S. *et al.* Air Pollution as a Risk Factor for Incident Chronic Obstructive Pulmonary Disease and Asthma. A 15-Year Population-based Cohort Study. *Am. J. Respir. Crit. Care Med.* **203**, 1138–1148 (2021).

112. Salvi, S. & Barnes, P. J. Is Exposure to Biomass Smoke the Biggest Risk Factor for COPD Globally? *Chest* **138**, 3–6 (2010).

113. European Respiratory Society. Occupational Risk Factors. in *European Lung White Book*

114. Govender, N., Lalloo, U. G. & Naidoo, R. N. Occupational Exposures and Chronic Obstructive Pulmonary Disease: A Hospital Based Case - Control Study. *Thorax* **66**, 597–601 (2011).

115. Viscardi, R. M. & Hasday, J. D. Role of Ureaplasma Species in Neonatal Chronic Lung Disease: Epidemiologic and Experimental Evidence. *Pediatr. Res.* **65**, 84–90 (2009).

116. Huang, Y. J. & Lynch, S. V. The Emerging Relationship between the Airway Microbiota and Chronic Respiratory Disease: Clinical Implications. *Expert Rev. Respir. Med.* **5**, 809–821 (2014).

117. Budden, K. F. *et al.* Functional Effects of the Microbiota in Chronic Respiratory Disease. *Lancet Respir. Med.* **7**, 907–920 (2019).

118. Hanson, C., Rutten, E. P., Wouters, E. F. M. & Rennard, S. Influence of Diet and Obesity on COPD Development and Outcomes. *Int. J. Chron. Obstruct. Pulmon. Dis.* **9**, 723–733 (2014).

119. Sullivan, K. & Thakur, N. Structural and Social Determinants of Health in Asthma in Developed Economies: a Scoping Review of Literature Published Between 2014 and 2019. *Curr. Allergy Asthma Rep.* **20**, (2020).

120. Kuhn, B. T., Wick, K. D. & Schivo, M. An Update in Health Disparities in COPD in the USA. *Curr. Pulmonol. Reports* **10**, 14–21 (2021).

121. Chen, Y. *et al.* Income Adequacy and Education Associated with the Prevalence of Obesity in Rural Saskatchewan, Canada. *BMC Public Health* **15**, 1–7 (2015).

122. Kim, T. J. & Knesebeck, O. von dem. Income and Obesity: What is the Direction of the Relationship? A Systematic Review and Meta-analysis. *Public Heal. Res.* **8**, e019862 (2018).

123. Goodman, E. & Huang, B. Socioeconomic Status, Depressive Symptoms, and Adolescent Substance Use. *Arch. Pediatr. Adolesc. Med.* **156**, 448–453 (2002).

124. Quon, E. C. & McGrath, J. J. Subjective Socioeconomic Status and Adolescent Health: A Meta-analysis. *Heal. Psychol.* **33**, 433 (2013).

125. Sahni, S., Talwar, A. A., Khanijo, S. & Talwar, A. A. Socioeconomic Status and its Relationship to Chronic Respiratory Disease. *Adv. Respir. Med.* **85**, 97–108 (2017).

126. Whitehead, M. & Dahlgren, G. *Concepts and Principles for Tackling Social Inequities in Health: Levelling up Part 1*. (2006).

127. Jin, Y., Zhu, D. & He, P. Social Causation or Social Selection? The Longitudinal Interrelationship Between Poverty and Depressive Symptoms in China. *Soc. Sci. Med.* **249**, 1–13 (2020).

128. Roux, A. V. D. *et al.* Neighborhood of Residence and Incidence of Coronary Heart Disease. *N. Engl. J. Med.* **345**, 99–106 (2009).

129. Yen, I. H. & Syme, S. L. The Social Environment and Health: A Discussion of the Epidemiologic Literature. **20**, 287–308 (1999).

130. Galobardes, B., Lynch, J. W. & Smith, G. D. Childhood Socioeconomic Circumstances and Cause-specific Mortality in Adulthood: Systematic Review and Interpretation. *Epidemiol. Rev.* **26**, 7–21 (2004).

131. West, P. Rethinking the Health Selection Explanation for Health Inequalities. *Soc. Sci. Med.* **32**, 373–384 (1991).

132. Uphoff, E. *et al.* A Systematic Review of Socioeconomic Position in Relation to Asthma and Allergic Diseases. *Eur. Respir. J.* **46**, 364–374 (2015).

133. Khullar, D. & Chokshi, D. A. Health, Income, and Poverty: Where We are and What could Help. *Health Aff.* **10**, (2018).

134. Williamson, D. L. *et al.* Low-income Canadians' Experiences with Health-related Services: Implications for Health Care Reform. *Health Policy (New. York).* **76**, 106–121 (2006).

135. Federico, B., Costa, G. & Kunst, A. E. Educational Inequalities in Initiation, Cessation, and Prevalence of Smoking Among 3 Italian Birth Cohorts. *Am. J. Public Health* **97**, 838–845 (2007).

136. Wellman, R. J. *et al.* Socioeconomic Status is Associated with the Prevalence and Co-occurrence of Risk Factors for Cigarette Smoking Initiation During Adolescence. *Int. J. Public Health* **63**, 125–136 (2018).

137. Bertozzi, S. M. & Gutiérrez, J. P. Poverty, Cash Transfers, and Risk Behaviours. *Lancet Glob. Heal.* **1**, e315–e316 (2013).

138. Bolland, J. M. Hopelessness and Risk Behaviour among Adolescents Living in High-poverty Inner-city Neighbourhoods. *J. Adolesc.* **26**, 145–158 (2003).

139. Grafova, I. B. Financial Strain and Smoking. *J. Fam. Econ. Issues* **32**, 327–340 (2011).

140. Assari, S. & Bazargan, M. Unequal Effects of Educational Attainment on Workplace Exposure to Second-Hand Smoke by Race and Ethnicity; Minorities' Diminished Returns in the National Health Interview Survey (NHIS). *J. Med. Res. Innov.* **3**, e000179 (2019).

141. Wnkleby, M. A., Jatulis, D. E., Frank, E., Fortmann, S. P. & Winkleby, M. A. Socioeconomic Status and Health: How Education, Income, and Occupation Contribute to Risk Factors for Cardiovascular Disease. *Am. J. Public Health* **82**, 816–820 (1992).

142. Ettner, S. L. & Grzywacz, J. G. Socioeconomic Status and Health Among Californians: An Examination of Multiple Pathways. *Am. J. Public Health* **93**, 441–444 (2011).

143. Conlon, B. A. *et al.* Home Environment Factors and Health Behaviors of Low-income, Overweight, and Obese Youth. *Am. J. Health Behav.* **43**, 420–436 (2019).

144. Fukuda, Y., Nakamura, K. & Takano, T. Accumulation of Health Risk Behaviours is Associated with Lower Socioeconomic Status and Women's Urban Residence: A Multilevel Analysis in Japan. *BMC Public Health* **5**, 1–10 (2005).

145. Renalds, A., Smith, T. H. & Hale, P. J. A Systematic Review of Built Environment and Health. *Fam. Community Heal.* **33**, 68–78 (2010).

146. The GOLD Team. Global strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Diseases: 2022 Report. (2021). Available at: https://goldcopd.org/wp-content/uploads/2021/11/GOLD-REPORT-2022-v1.1-22Nov2021_WMV.pdf. (Accessed: 4th December 2021)

147. Cai, L. *et al.* Socioeconomic Variations in Chronic obstructive Pulmonary Disease Prevalence, Diagnosis, and Treatment in Rural Southwest China. *BMC Public Health* **20**, 536 (2020).

148. Yin, P., Zhang, M., Li, Y., Jiang, Y. & Zhao, W. Prevalence of COPD and its Association with socioeconomic status in China: Findings from China Chronic Disease Risk Factor Surveillance 2007. *BMC Public Health* **11**, 586 (2011).

149. Levin, K. A., Anderson, D. & Crighton, E. Prevalence of COPD by Age, Sex, Socioeconomic Position and Smoking status; A Cross-sectional Study. *Health Educ.* **120**, 275–288 (2020).

150. Pahwa, P. *et al.* Incidence and Longitudinal Changes in Prevalence of Chronic Bronchitis in Farm and Non-Farm Rural Residents of Saskatchewan. *J. Occup. Environ. Med.* **61**, 347–356 (2019).

151. Coogan, P. F. *et al.* Neighborhood and Individual Socioeconomic Status and Asthma Incidence in African American Women. *Ethn. Dis.* **26**, 113–122 (2016).

152. Ejlskov, L. *et al.* The Effect of Early-life and Adult Socioeconomic Position on Development of Lifestyle-related Diseases. *Eur. J. Public Health* **29**, 562–567 (2019).

153. Pahwa, P. *et al.* The Saskatchewan Rural Health Study: An Application of a Population Health Framework to Understand Respiratory Health Outcomes. *BMC Res. Notes* **5**, 1–13 (2012).

154. Pahwa, P. *et al.* Cohort profile: The Saskatchewan Rural Health Study - Adult component. *BMC Res. Notes* **10**, 1–7 (2017).

155. Advisory Committee on Population Health. *Strategies for Population Health: Investing in the Health of Canadians*. (1994).

156. du Plessis, V., Beshiri, R., Bollman, D. R. & Clemenson, H. *Definition of 'Rural'. Agriculture and Rural Working Paper Series* (2002).

157. Dillman, D. A. *Mail and Internet Surveys: The Tailored Design Method*. (John Wiley & Sons Inc, 2007).

158. Dziak, J. J., Lanza, S. T. & Tan, X. Effect Size, Statistical Power, and Sample Size Requirements for the Bootstrap Likelihood Ratio Test in Latent Class Analysis. *Struct. Equ. Model. A Multidiscip. J.* **21**, 534–552 (2014).

159. Szanton, S. L. *et al.* Effect of Financial Strain on Mortality in Community-Dwelling Older Women. *J Gerontol B Psychol Sci Soc Sci.* **63**, s369–s374 (2008).

160. Mcniven, C., Puderer, H. & Janes, D. *Census Metropolitan Area and Census Agglomeration Influenced Zones (MIZ): A Description of the Methodology*. (2000).

161. Melki, I. S., Beydoun, H. A., Khogali, M., Tamim, H. & Yunis, K. A. Household Crowding Index: a Correlate of Socioeconomic Status and Inter-pregnancy Spacing in an Urban Setting. *J. Epidemiol. Community Heal.* **58**, 476–480 (2004).

162. Dahoo, I., Martin, W. & Stryhn, H. *Methods in Epidemiological Research*. (VER Inc, 2012).

163. Saegusa, T., Di, C. & Chen, Y. Q. Hypothesis Testing for an Extended Cox Model with Time-varying Coefficients. *Biometrics* **70**, 619–628 (2014).

164. Schoenfeld, D. Partial Residuals for The Proportional Hazards Regression Model. *Biometrika* **69**, 239–241 (1982).

165. UCLA Institute for Digital Research & Education. Survival Analysis with Stata. *Statistical Consulting* Available at: https://stats.idre.ucla.edu/stata/seminars/stata-survival/#fit. (Accessed: 7th July 2021)

166. Harrell, F. E., Califf, R. M., Pryor, D. B., Lee, K. L. & Rosati, R. A. Evaluating the Yield of Medical Tests. *JAMA* **247**, 2543–2546 (1982).

167. Kang, H. The Prevention and Handling of the Missing Data. *Korean J. Anesthesiol.* **64**, 406 (2013).

168. Olinsky, A., Chen, S. & Harlow, L. The Comparative Efficacy of Imputation Methods for Missing Data in Structural Equation Modeling. *Eur. J. Oper. Res.* **151**, 53–79 (2003).

169. Barnard, J. & Meng, X. L. Applications of Multiple Imputation in Medical Studies: From

AIDS to NHANES. *Stat. Methods Med. Res.* **8**, 17–36 (1999).

170. Allison, P. Listwise Deletion: It's NOT Evil . *Statistical Horizons* (2014). Available at: https://statisticalhorizons.com/listwise-deletion-its-not-evil/. (Accessed: 20th May 2022)

171. Nielsen, M., Bårnes, C. B. & Ulrik, C. S. Clinical Characteristics of the Asthma–COPD Overlap Syndrome – A Systematic Review. *Int. J. Chron. Obstruct. Pulmon. Dis.* **10**, 1454 (2015).

172. Postma, D. S. & Rabe, K. F. The Asthma-COPD Overlap Syndrome. *N. Engl. J. Med.* **373**, 1241–1249 (2015).

173. Bosonea, A. M. *et al.* Developments in asthma incidence and prevalence in Alberta between 1995 and 2015. *Allergy, Asthma Clin. Immunol.* **16**, 1–11 (2020).

174. Stúrmer, T. & Brookhart, A. Study Design Considerations. in *Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide* (eds. Valentgas, P., Dreyer, N., Nourjah, P., Smith, S. & Torchia, M.) **Publication No. 12**, 21–33 (MD: Agency for Healthcare Research and Quality, 2013).

175. Chen, Y., Dales, R., Tang, M. & Krewski, D. Association between Income Adequacy and Asthma in Canadians. in *Proceedings of Statistics Canada Symposium 2002: Modelling Survey Data for Social and Economic Research* (2002).

176. Hoffmann, R., Kröger, H. & Pakpahan, E. Pathways Between Socioeconomic Status and Health: Does Health Selection or Social Causation Dominate in Europe? *Adv. Life Course Res.* **36**, 23–36 (2018).

177. Lynch, J. & Kaplan, G. Social Position. in *Social Epidemiology* (eds. Berkman, L. & Kawachi I) 13–35 (Oxford University Press, 2000).

178. Lawlor, D. A., Smith, G. D. & Ebrahim, S. Socioeconomic Position and Hormone Replacement Therapy Use: Explaining the Discrepancy in Evidence from Observational and Randomized Controlled Trials. *Am. J. Public Health* **94**, 2149–2154 (2004).

179. Bjuggren, P. O. & Sund, L. G. A Transition Cost Rationale for Transition of the Firm within the Family. *Small Bus. Econ.* **19**, 123–133 (2002).

180. Case, A., Lubotsky, D. & Paxson, C. Economic Status and Health in Childhood: The Origins of the Gradient. *Am. Econ. Rev.* **92**, 1308–1334 (2002).

181. Deaton, A. Policy Implications of the Gradient of Health and Wealth. *Health Aff.* **21**, 13–30 (2002).

182. Lenhart, O. The Effects of Income on Health: New Evidence from the Earned Income Tax Credit. *Rev. Econ. Househ.* **17**, 377–410 (2019).

183. Beaglehole, R. *et al.* Priority Actions for the Non-communicable Disease Crisis. *Lancet* **377**, 1438–1447 (2011).

184. Selye, H. *The Stress of Life*. (McGraw Hill, 1956).

185. Kiecolt-Glaser, J. K., McGuire, L., Robles, T. F. & Glaser, R. Psychoneuroimmunology:

Psychological Influences on Immune Function and Health. *J. Consult. Clin. Psychol.* **70**, 537–547 (2002).

186. Lehrer, P., Feldman, J., Giardino, N., Song, H. S. & Schmaling, K. Psychological Aspects of Asthma. *J. Consult. Clin. Psychol.* **70**, 691–711 (2002).

187. Cohen, S. *et al.* Reactivity and Vulnerability to Stress-Associated Risk for Upper Respiratory Illness. *Psychosom. Med.* **64**, 302–310 (2002).

188. Woolf, S. H. *et al. How are Income and Wealth Linked to Health And Longevity?* (2015).

189. Wen, M., Browning, C. R. & Cagney, K. A. Poverty, Affluence, and Income Inequality: Neighborhood Economic Structure and its Implications for Health. *Soc. Sci. Med.* **57**, 843–860 (2003).

190. Peters, D. H. *et al.* Poverty and Access to Health Care in Developing Countries. *Ann. N. Y. Acad. Sci.* **1136**, 161–171 (2008).

191. Schneider, M., Bradshaw, D., Norman, R., Steyn, K. & Laubscher, R. Poverty and Non-communicable Diseases in South Africa. *Scand. J. Public Health* **37**, 176–186 (2009).

192. Li, Y. *et al.* Factors Affecting Catastrophic Health Expenditure and Impoverishment from Medical Expenses in China: Policy Implications of Universal Health Insurance. *Bull World Heal. Organ* **90**, 664–671 (2012).

193. Xu, K. *et al.* Household Catastrophic Health Expenditure: A Multicountry Analysis. *Lancet* **362**, 111–117 (2003).

194. Zhou, Y. *et al.* The Association between BMI and COPD: The Results of Two Population-based Studies in Guangzhou, China. *COPD J. Chronic Obstr. Pulm. Dis.* **10**, 567–572 (2013).

195. Hu, S. *et al.* Reform of How Health Care is Paid for in China: Challenges and Opportunities. *Lancet* **372**, 1846–1853 (2008).

196. Bonds, M. H., Keenan, D. C., Rohani, P. & Sachs, J. D. Poverty Trap Formed by the Ecology of Infectious Diseases. *Proc. R. Soc. B* **277**, 1185–1192 (2010).

197. Gaffney, A. W. *et al.* Health Care Disparities Across the Urban-Rural Divide: A National Study of Individuals with COPD. *J. Rural Heal.* (2020). doi:10.1111/JRH.12525

198. Rabinowitz, P. M. *et al.* Proximity to Natural Gas Wells and Reported Health Status: Results of a Household Survey in Washington County, Pennsylvania. *Environ. Health Perspect.* **123**, 21–26 (2015).

199. Coogan, P. F. *et al.* Active and Passive Smoking and the Incidence of Asthma in the Black Women's Health Study. *Am. J. Respir. Crit. Care Med.* **191**, 168–176 (2015).

200. Hedlund, U., Eriksson, K. & Rönmark, E. Socio-economic Status is Related to Incidence of Asthma and Respiratory Symptoms in Adults. *Eur. Respir. J.* **28**, 303–310 (2006).

201. Torén, K. *et al.* A Prospective study of Asthma Incidence and its Predictors: The RHINE Study. *Eur. Respir. J.* **24**, 942–946 (2004).

202. Lindberg, A. *et al.* Seven-Year Cumulative Incidence of COPD in an Age-Stratified General Population Sample. *Chest* **129**, 879–885 (2006).

203. Kojima, S. *et al.* Incidence of Chronic Obstructive Pulmonary Disease, and the Relationship between Age and Smoking in a Japanese Population. *J. Epidemiol.* **17**, 54–60 (2007).

204. Forey, B. A., Thornton, A. J. & Lee, P. N. Systematic Review with Meta-analysis of the Epidemiological Evidence Relating Smoking to COPD, Chronic Bronchitis and Emphysema. *BMC Pulm. Med.* **11**, 1–61 (2011).

205. Hersh, C. P. *et al.* Family History Is a Risk Factor for COPD. *Chest* **140**, 343–350 (2011).

206. Bui, D. S. *et al.* Trajectories of Asthma and Allergies from 7 Years to 53 Years and Associations with Lung Function and Extrapulmonary Comorbidity Profiles: A Prospective Cohort Study. *Lancet Respir. Med.* **9**, 387–396 (2021).

207. Corry, D. B. & Kheradmand, F. Toward a Comprehensive Understanding of Allergic Lung Disease. *Trans. Am. Clin. Climatol. Assoc.* **120**, 48 (2009).

208. Lommatzsch, M. *et al.* IgE is Associated with Exacerbations and Lung Function Decline in COPD. *Respir. Res.* **23**, 1–9 (2022).

209. Ehrlich, S. F., Quesenberry, C. P., Van Den Eeden, S. K., Shan, J. & Ferrara, A. Patients Diagnosed With Diabetes Are at Increased Risk for Asthma, Chronic Obstructive Pulmonary Disease, Pulmonary Fibrosis, and Pneumonia but Not Lung Cancer. *Diabetes Care* **33**, (2010).

210. Thomsen, S. F., Duffy, D. L., Kyvik, K. O., Skytthe, A. & Backer, V. Risk of Asthma in Adult Twins with Type 2 Diabetes and Increased Body Mass Index. *Allergy* **66**, 562–568 (2011).

211. Fuller-Thomson, E., Howden, K. E. N., Fuller-Thomson, L. R. & Agbeyaka, S. A Strong Graded Relationship between Level of Obesity and COPD: Findings from a National Population-Based Study of Lifelong Nonsmokers. *J. Obes.* **2018**, (2018).

212. Barnes, A. S. & Coulter, S. A. The Epidemic of Obesity and Diabetes: Trends and Treatments. *Texas Hear. Inst. J.* **38**, 144 (2011).

213. Khateeb, J., Fuchs, E. & Khamaisi, M. Diabetes and Lung Disease: An Underestimated Relationship. *Rev. Diabet. Stud.* **15**, 15 (2019).

214. Spectrum Health Lakeland Diabetes and Endocrinology. Diabetes Health Library. Available at: https://www.spectrumhealthlakeland.org/lakeland-diabetes/diabetes-health-library/Content/3/60033. (Accessed: 20th September 2022)

215. Konrad, S., Hossain, ; A, Senthilselvan, ; A, Dosman, J. A. & Pahwa, ; P. Chronic Bronchitis in Aboriginal People-Prevalence and Associated Factors. *Chronic Dis. Inj. Can.* **33**, 218–225 (2013).

216. Fedak, K. M., Bernal, A., Capshaw, Z. A. & Gross, S. Applying the Bradford Hill Criteria in the 21st Century: How Data Integration has Changed Causal Inference in Molecular

Epidemiology. *Emerg. Themes Epidemiol.* **12**, 14 (2015).

217. Bethea, T. N., Lopez, R. P., Cozier, Y. C., White, L. F. & Mcclean, M. D. The Relationship Between Rural Status, Individual Characteristics, and Self-Rated Health in the Behavioral Risk Factor Surveillance System. *J. Rural Heal.* **28**, 327–338 (2012).

218. Pampalon, R., Hamel, D. & Gamache, P. Health Inequalities in Urban and Rural Canada: Comparing Inequalities in Survival According to an Individual and Area-Based Deprivation Index . *Health Place* **16**, 416–420 (2010).

219. Yawn, B. P. Differential Assessment and Management of Asthma vs Chronic Obstructive Pulmonary Disease. *Medscape J. Med.* **11**, 20 (2009).

220. Assari, S., Chalian, H. & Bazargan, M. Race, Ethnicity, Socioeconomic Status, and Chronic Lung Disease in the U.S. *Res. Heal. Sci.* **5**, p48 (2020).

221. Martin, L. M., Leff, M., Calonge, N., Garrett, C. & Nelson, D. E. Validation of Self-reported Chronic Conditions and Health Services in a Managed Care Population. *Am. J. Prev. Med.* **18**, 215–218 (2000).

222. Najafi, F., Moradinazar, M., Hamzeh, B. & Rezaeian, S. The Reliability of Self-reporting Chronic Diseases: How Reliable is the Result of Population-based Cohort Studies. *J. Prev. Med. Hyg.* **60**, E353 (2019).

223. Ekholm, O., Gundgaard, J., Hansen, E. H. & Rasmussen, N. K. r. The Effect of Health, Socio-economic Position, and Mode of Data Collection on Non-response in Health Interview Surveys. *Scand. J. Public Health* **38**, 699–706 (2010).

224. Government of Saskatchewan. Covered Population 2014. *Health Information Solutions Centre, Saskatchewan Ministry of Health.* 156 (2014). Available at: http://population.health.gov.sk.ca/CovPopBook2014.pdf. (Accessed: 12th July 2022)

225. National Rural Health Association. National Health Policy Reform: The Rural Perspective. (1992). Available at: https://www.ruralhealth.us/getattachment/Advocate/Policy-Documents-(ME)/Archive-(ME)-(1)/RuralPerspective.pdf.aspx?lang=en-US. (Accessed: 7th July 2022)

226. National Rural Health Policy Forum. *Healthy horizons : A Framework for Improving the Health of Rural, Regional and Remote Australians. Canberra : National Rural Health Policy Forum and the National Rural Health Alliance for the Australian Health Ministers' Conference, 1999* (National Rural Health Policy Forum and the National Rural Health Alliance for the Australian Health Ministers' Conference, 1999).

227. Cawley, J. F. & Hooker, R. S. Physician Assistants in American Medicine: The Half-Century Mark. *Am. J. Manag. Care* **19**, e333–e341 (2013).

228. British Columbia Ministry of Health. *Rural Health Services in BC: A Policy Framework to Provide a System of Quality Care*. (2015).

229. Wilson, R., Rourke, J., Oandasan, I. F. & Bosco, C. Progress Made on Access to Rural Health Care in Canada. *Can. J. Rural Med.* **66**, 31–36 (2020).

230. Carr, D. Improving the Health of the World's Poorest People. in *The Health Bulletin* (Population Reference Bureau, 2004).

APPENDIX

A.1. ETHICS

A.1.1. Ethics Approval

UNIVERSITY OF SASKATCHEWAN

Biomedical Research Ethics Board (Bio-REB) 09-Sep-2021

Certificate of Approval

Application ID: 2902

Principal Investigator: Bonnie Janzen

Department: Department of Community Health and Epidemiology

Locations Where Research Activities are Conducted: Data collection is complete. Data will be analyzed in Saskatoon., Canada

Student(s): Francis Abayateye

Funder(s):

Sponsor: Canadian Institutes of Health Research

Title: Incidence of Chronic Obstructive Lung Disease and Longitudinal Association with Socioeconomic Status (SES) among Adults in Rural Saskatchewan

Protocol Number:

Approved On: 07-Sep-2021

Expiry Date: 07-Sep-2022

Approval Of:
* Ethics Application_v3
* Variable List

Acknowledgment Of:
* Notice of Ethical Review Response
* McMaster Chart Review Tutorial Certificate of Completion for Francis Abayateye
* TCPS2 CoreTutorial Certificate of Completion for Francis Abayateye
* Reviewed with COVID-19 safety considerations in mind

Review Type: Delegated Review

IRB Registration Number: Not Applicable

1 / 2

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

A.1.2. Ethics Re-approval

UNIVERSITY OF SASKATCHEWAN

Biomedical Research Ethics Board (Bio-REB) 08-Sep-2022

Certificate of Re-Approval

Application ID: 2902

Principal Investigator: Bonnie Janzen

Department: Department of Community Health and Epidemiology

Locations Where Research Activities are Conducted: Data collection is complete. Data will be analyzed in Saskatoon., Canada

Student(s): Francis Abayateye

Funder(s):

Sponsor: Canadian Institutes of Health Research

Title: Incidence of Chronic Obstructive Lung Disease and Longitudinal Association with Socioeconomic Status (SES) among Adults in Rural Saskatchewan

Approval Effective Date: 08-Sep-2022

Expiry Date: 08-Sep-2023

Acknowledgment Of:

* Certificate of completion for The TCPS 2 Francis Abayateye
* McMaster University Chart Review Tutorial Francis Abayateye

Review Type: Delegated Review

IRB Registration Number: Not Applicable

* This study, inclusive of all previously approved documents, has been re-approved until the expiry date noted above

1 / 2

A.2. STATA CODES

```
**THESIS DO FILE - LONGITUDINAL ASSOCIATION OF SES WITH CRC**
clear
set more off
cd "C:\Users\sir16\iCloudDrive\Desktop\2020 & BEYOND\CH&E\THESIS\Data
Analysis\Data"
use "SRHS_Longitudinal data"

****Selecting Variables of interest and renaming****
keep HOUSEID PERSONID MIZ TOWN_RM MUNICIPALITY QUADRANT KEY
ReceiveDt2 h_DATE FOLLOWUPIDN BOTHSTUDIES ///
EDUGRP EDUGRP2 INCOMEADQR INCOMEADQR2 INCOMEADQ INCOMEADQ2
INCOMEADQR INCOMEADQR2 INCOMEGRP INCOMEGRP2 h_EXTRAMONEY ///
h_EXTRAMONEY2 AGEGRP AGEGRP2 i_SEX i_SEX2 MARITALGRP MARITALGRP2
BMIGRP BMIGRP2 FARM FARM2 h_PEOPLE h_PEOPLE2 h_COAL ///
h_COAL2 SMKSTAT SMKSTAT2 i_ALCOHOL i_ALCOHOL2 EXERCISELONG
EXERCISELONG2 i_FARM1STYR i_FARM1STYR2 i_MOMSMOKE BEDRMGRP ///
i_SMOKESTUBBLE i_SMOKESTUBBLE2 i_DIESELFUMES i_DIESELFUMES2
ri_WELDING i_WELDING2 ri_SOLVENT i_SOLVENT2 ///
ri_RADIATION i_RADIATION2 ri_WOODDUST i_WOODDUST2 ri_GRAINDUST
i_GRAINDUSTEXP2 ri_ASBESTOSDUST i_ASBESTOSDUST2 ///
ri_MINEDUSTEXP i_MINEDUSTEXP2 ri_DADLUNGDIS ri_MOMLUNGDIS
ri_SIBLUNGDIS ROUTINEDISTGRP ri_HEARTDIS i_HEARTDIS2 ///
ri_OILGAS i_OILGAS2 ri_HERBICIDES i_HERBICIDES2 ri_FUNGICIDES
i_FUNGICIDES2 ri_INSECT i_INSECT2 ri_MOLDS i_MOLDS2 i_HEARTATT ///
i_HEARTATT2 ri_EMPHYSEVER i_EMPHYSEVER2 ri_CHRBRONEVER
i_CHRBRONEVER2 ri_COPDEVER i_COPDEVER2 EVERASTHMADR
EVERASTHMADR2 ///
h_BEDROOMS h_BEDROOMS2 h_NATGAS h_NATGAS2 h_DAMP h_DAMP2
rh_CIGARETTES h_CIGARETTES2 HOUSEHOLDSMK HOUSEHOLDSMK2 ///
h_SMELL h_SMELL2 ONEPARENTLUNG i_ETHNICBG i_AGE i_AGE2
ri_ALLERGYHOUSEDUST i_ALLERGYHOUSEDUST2 ri_ALLERGYCAT
i_ALLERGYCAT2 ///
ri_ALLERGYDOGS i_ALLERGYDOG2 ri_ALLERGYGRASSES i_ALLERGYGRASSES2
ri_ALLERGYPOLLENS i_ALLERGYPOLLENS2 ri_ALLERGYMOLDS ///
i_ALLERGYMOLDS2 ri_OTHERALLERGY i_OTHERALLERGY2 ANYALLERGY
ANYALLERGY2 i_ETHNICBG ri_DIABETES i_DIABETES2 h_HOUSEINCOME ///
i_GRAINHOWOFTEN i_MINEHOWOFTEN i_ASBESHOWOFTEN i_WOODHOWOFTEN
i_SMOKEHOWOFTEN i_DIESELHOWOFTEN i_WELDINGHOWOFTEN ///
i_SOLVENTHOWOFTEN i_OILGASHOWOFTEN i_HERBHOWOFTEN
i_FUNGHOWOFTEN i_INSECTHOWOFTEN i_MOLDHOWOFTEN
i_RADIATIONHOWOFTEN ///
h_HOUSEINCOME2 i_EDUCATION i_EDUCATION2 i_AGE
```

```
***Renaming Variables***
rename ri_WELDING i_WELDING
rename ri_SOLVENT i_SOLVENT
rename ri_RADIATION i_RADIATION
rename ri_WOODDUST i_WOODDUST
rename ri_GRAINDUST i_GRAINDUST
rename ri_ASBESTOSDUST i_ASBESTOSDUST
rename ri_MINEDUSTEXP i_MINEDUSTEXP
rename ri_HEARTDIS i_HEARTDIS
rename ri_OILGAS i_OILGAS
rename ri_HERBICIDES i_HERBICIDES
rename ri_FUNGICIDES i_FUNGICIDES
rename ri_INSECT i_INSECT
rename ri_MOLDS i_MOLDS
rename ri_EMPHYSEVER i_EMPHYSEVER
rename ri_CHRBRONEVER i_CHRBRONEVER
rename ri_COPDEVER i_COPDEVER
rename rh_CIGARETTES h_CIGARETTES
rename i_GRAINDUSTEXP2 i_GRAINDUST2
rename ri_MOMLUNGDIS i_MOMLUNGDIS
rename ri_DADLUNGDIS i_DADLUNGDIS
rename ri_SIBLUNGDIS i_SIBLUNGDIS
rename ri_ALLERGYHOUSEDUST i_ALLERGYHOUSEDUST
rename ri_ALLERGYCAT i_ALLERGYCAT
rename ri_ALLERGYDOGS i_ALLERGYDOG
rename ri_ALLERGYGRASSES i_ALLERGYGRASSES
rename ri_ALLERGYPOLLENS i_ALLERGYPOLLENS
rename ri_ALLERGYMOLDS i_ALLERGYMOLDS
rename ri_OTHERALLERGY i_OTHERALLERGY
rename ri_DIABETES i_DIABETES


****Renaming varibles for shape tranformation****
foreach var of varlist i_SEX i_SMOKESTUBBLE i_DIESELFUMES i_WELDING
i_SOLVENT i_OILGAS i_HERBICIDES ANYALLERGY h_COAL ///
i_FUNGICIDES i_INSECT i_MOLDS EDUGRP INCOMEGRP INCOMEADQ i_AGE
AGEGRP MARITALGRP BMIGRP h_EXTRAMONEY FARM SMKSTAT ///
i_ALCOHOL EXERCISELONG i_FARM1STYR i_EMPHYSEVER i_CHRBRONEVER
i_COPDEVER EVERASTHMADR i_HEARTDIS i_DIABETES h_BEDROOMS ///
h_PEOPLE h_NATGAS h_DAMP h_CIGARETTES h_SMELL i_RADIATION
i_WOODDUST i_GRAINDUST i_ASBESTOSDUST i_MINEDUSTEXP INCOMEADQR ///
 i_ALLERGYHOUSEDUST i_ALLERGYCAT i_ALLERGYDOG i_ALLERGYGRASSES
i_ALLERGYPOLLENS i_ALLERGYMOLDS i_OTHERALLERGY i_HEARTATT ///
 i_EDUCATION h_HOUSEINCOME {

rename `var' `var'1
```

```
}

***Reshaping from wide to long***
reshape long i_SEX BMIGRP  FARM  INCOMEADQ INCOMEADQR INCOMEGRP  i_AGE
AGEGRP EDUGRP MARITALGRP i_FARM1STYR SMKSTAT ///
i_ALCOHOL  i_HEARTDIS  EXERCISELONG EVERASTHMADR i_EMPHYSEVER
i_CHRBRONEVER i_COPDEVER i_SMOKESTUBBLE i_HEARTATT ///
i_DIESELFUMES i_WELDING i_SOLVENT i_OILGAS i_HERBICIDES i_FUNGICIDES
i_INSECT i_MOLDS h_EXTRAMONEY h_COAL ///
h_PEOPLE h_BEDROOMS h_NATGAS h_DAMP h_CIGARETTES h_SMELL
i_RADIATION i_WOODDUST i_GRAINDUST i_ASBESTOSDUST ///
i_MINEDUSTEXP i_ALLERGYHOUSEDUST i_ALLERGYCAT i_ALLERGYDOG
i_ALLERGYGRASSES i_ALLERGYPOLLENS i_ALLERGYMOLDS ///
i_OTHERALLERGY ANYALLERGY i_DIABETES i_EDUCATION h_HOUSEINCOME,
i(HOUSEID PERSONID) j(time)


***Generating CRC variable***

**COPD**
gen copd = 1 if i_COPDEVER ==1| i_CHRBRONEVER == 1| i_EMPHYSEVER == 1
replace copd = 0 if i_COPDEVER ==2 & i_CHRBRONEVER == 2 & i_EMPHYSEVER == 2
label variable copd "Has a doctor ever said you have COPD?"
label define copd 1 "Yes" 0 "No"
label value copd copd


**Asthma**
gen asthma = 1 if EVERASTHMADR == 1
replace asthma = 0 if EVERASTHMADR == 2
label variable asthma "Asthma"
label define asthma 0 "Not developed" 1 "Developed"
label value asthma asthma

**CRC**
gen CRC = 1 if copd == 1| asthma == 1
replace CRC = 0 if copd == 0 & asthma == 0
label variable CRC "Have Chronic Respiratory Conditions"
label define CRC 1 "developed" 0 "not developed"
label value CRC CRC


gen ACOS = 1 if copd ==1 & asthma == 1
replace ACOS = 0 if copd == 0| asthma == 0 |(copd == 0& asthma == 0)
label variable ACOS "Asthma COPD Overlap Syndrome"
label define ACOS 1 "developed" 0 "not developed"
```

```
label value ACOS ACOS

****Generating new variables and cleaning data****
//Quadrant//
gen quad = 1 if QUADRANT ==1
replace quad = 2 if QUADRANT == 2
replace quad = 3 if QUADRANT == 3
replace quad = 4 if QUADRANT == 4
label var quad "Four corners"
label define quad 1 "SW" 2 "SE" 3 "NE" 4 "NW"
label value quad quad

//MIZ//
gen miz = 1 if MIZ ==1
replace miz = 2 if MIZ ==2
replace miz = 3 if MIZ ==3
replace miz = . if MIZ ==8
label variable miz "Metropolitan Influence Zone"
label define miz 1 "No MIZ" 2 "Weak MIZ" 3 "Moderate MIZ"
label value miz miz


//Number of Bedrooms//
gen bedrooms = 0 if BEDRMGRP == 1
replace bedrooms = 1 if BEDRMGRP == 2
replace bedrooms = . if BEDRMGRP == 8
label variable bedrooms "number of bedrooms in house"
label define bedrooms 0 "3 or less bedrooms" 1 "more than 3 bedrooms"
label value bedrooms bedrooms

//Household density//
gen hhsize = 1 if h_PEOPLE <=2
replace hhsize = 2 if h_PEOPLE > 2
replace hhsize = . if h_PEOPLE == 88
label var hhsize "Number of people in household"
label define hhsize 1 "2 or less" 2 "More than 2"
label value hhsize hhsize

//Crowding index//
gen c_index = h_PEOPLE/h_BEDROOMS
gen hc_index = 1 if c_index < 1
replace hc_index =2 if c_index >=1 & c_index <= 2
replace hc_index = 3 if c_index > 2
label variable hc_index "Household crowding index"
label define hc_index 1 "< 1" 2 "1 - 2" 3 "> 2"
label val hc_index hc_index
```

```
//Alcohol consumption//
gen alcohol = 1 if i_ALCOHOL == 1
replace alcohol = 2 if i_ALCOHOL == 2| i_ALCOHOL == 3
replace alcohol = 3 if i_ALCOHOL == 4
replace alcohol = 4 if i_ALCOHOL == 5| i_ALCOHOL == 6
replace alcohol = 5 if i_ALCOHOL == 7| i_ALCOHOL == 8
replace alcohol = . if i_ALCOHOL == 88
label var alcohol "Alcohol consumption in the past 12 month"
label define alcohol 1 "never" 2 "1 or less a month" 3 "2-3times a month" ///
4 "1-3times a week" 5 "4 or more a week"
label value alcohol alcohol
```

****There are other categories: don't know & multiple response as categories***

```
//First year on Farm// /
gen farm1styr =  1 if i_FARM1STYR == 1
replace farm1styr = 0 if i_FARM1STYR == 2
label var farm1styr "lived on farm during first year of life"
label define farm1styr 1 "Yes" 0 "No"
label value farm1styr farm1styr

gen momsmoke = i_MOMSMOKE
replace momsmoke =. if i_MOMSMOKE == 8 |i_MOMSMOKE ==7
label var momsmoke "Did your mother smoke when she was pregnant with you?"
label define momsmoke 1 "Yes" 2 "No"
label value momsmoke momsmoke

// One parent had lung disease //
gen oneparentlung = 1 if ONEPARENTLUNG == 1
replace oneparentlung = 0 if ONEPARENTLUNG == 0
label var oneparentlung "one parent had lung disease"
label define oneparentlung 1 "Yes" 0 "No"
label value oneparentlung oneparentlung


//Cigarette smoking at home//
gen h_cigarette = 1 if HOUSEHOLDSMK == 1
replace h_cigarette = 0 if HOUSEHOLDSMK == 2
label variable h_cigarette "People at home use cigarette at home"
label define h_cigarette 1 "Yes" 0 "No"
label value h_cigarette h_cigarette

//Mildew odor or musty smell//
gen smell = 1 if h_SMELL ==1
replace smell = 0 if h_SMELL ==2
```

```
label variable smell "home frequently has mildew odor or musty smell"
label define smell 1 "Yes" 0 "No"
label value smell smell

//Heart disease//
gen heartdis = 1 if i_HEARTDIS ==1
replace heartdis = 0 if i_HEARTDIS ==2
label variable heartdis "Dr ever said you have heart disease"
label define heartdis 1 "Yes" 0 "No"
label value heartdis heartdis

gen heartatt = 1 if i_HEARTATT == 1
replace heartatt = 0 if i_HEARTATT ==2
replace heartatt =. if i_HEARTATT==7| i_HEARTATT==8
label variable heartatt "Heart attack"
label define heartatt 1 "Yes" 0 "No"
label value heartatt heartatt

***These variables have don't know as a categories: (NO multiple response) ****
//Household Dampness//
gen damp = 1 if h_DAMP ==1
replace damp = 0 if h_DAMP ==2
label variable damp "water or dampness at home in past 12 month"
label define damp 1 "Yes" 0 "No"
label value damp damp

//Household Natural Gas usage//
gen natgas = 1 if h_NATGAS == 1
replace natgas = 0 if h_NATGAS == 2
label variable natgas "natural gas"
label define natgas 1 "Yes" 0 "No"
label value natgas natgas

***Missing values in these variables are either coded as 8 or 888***
//Recoding missing values//
gen edugrp = EDUGRP
replace edugrp =. if EDUGRP ==8
label var edugrp "highest educational level"
label define edugrp 1 "=< Grade 12" 2 "> Grade 12"
label value edugrp edugrp

gen edu3gp = 1 if i_EDUCATION == 1
replace edu3gp = 2 if i_EDUCATION == 2
replace edu3gp = 3 if i_EDUCATION == 3|i_EDUCATION==4
replace edu3gp = . if i_EDUCATION == 8
label var edu3gp "Three category educational level"
```

```
label define edu3gp 1 "Less than high school" 2 "High school" 3 "Postsecondary"
label value edu3gp edu3gp

gen extramoney = h_EXTRAMONEY
replace extramoney =. if h_EXTRAMONEY ==8
label var extramoney "At the end of the month, how much do you have left?"
label define extramoney 1 "Not enough money" 2 "Just enough money" 3 "Some money"
label value extramoney extramoney

gen income = INCOMEGRP
replace income =. if INCOMEGRP ==8
label var income "household income categories"
label define income 1 "< $20,000" 2 "$20,000- $39,999" 3 "$40,000- $59,999" 4 ">=$60,000"
label value income income

gen incomeadq = INCOMEADQ
replace incomeadq =. if INCOMEADQ == 8
label var incomeadq "Income adequacy"
label define incomeadq 1 "Lowest income" 2 "Lowest middle income" 3 "Upper middle income"
4 "Highest income"
label value incomeadq incomeadq

gen inc_new = 1 if h_HOUSEINCOME <=3
replace inc_new = 2 if h_HOUSEINCOME ==4|h_HOUSEINCOME==5
replace inc_new = 3 if h_HOUSEINCOME ==6|h_HOUSEINCOME==7
replace inc_new = 4 if h_HOUSEINCOME ==8
replace inc_new = . if h_HOUSEINCOME ==88
label var inc_new "Household income categories"
label define inc_new 1 "<=$29,999" 2 "$30,000 - $49,999" 3 "$50,000 - $79,999" 4 ">=
$80,000"
label value inc_new inc_new

gen income_ru = INCOMEADQR
replace income_ru =. if INCOMEADQR == 8
label var income_ru "Rural income"
label define income_ru 1 "Low income" 2 "High income"
label value income_ru income_ru

gen agegrp = AGEGRP
replace agegrp =. if AGEGRP == 8
label var agegrp "age group"
label define agegrp 1 "18-45 years" 2 "46-55 years" 3 "56-65 years" 4 "> 65 years"
label value agegrp agegrp

gen maritalgrp = 1 if MARITALGRP == 1
replace maritalgrp = 0 if MARITALGRP == 2
```

```
replace maritalgrp =. if MARITALGRP ==8
label var maritalgrp "marital status group"
label define maritalgrp 1 "Married/common-law/ living together" 0
"widowed/divorced/separated/never-married"
label value maritalgrp maritalgrp

gen sex = i_SEX
replace sex =. if i_SEX == 8
label var sex "Sex"
label define sex 1 "Male" 2 "Female"
label value sex sex

gen bmi = BMIGRP
replace bmi =. if BMIGRP ==8
label var bmi "Body Mass Index"
label define bmi 1 "Normal (< 25)" 2 "Overweight (>=25 and <=30)" 3 "Obese (>30)"
label value bmi bmi

gen farm = FARM
replace farm =. if FARM ==8
label var farm "Location of residence"
label define farm 1 "farm" 2 "Non-farm"
label value farm farm

gen smkstat = SMKSTAT
replace smkstat =. if SMKSTAT ==8
label var smkstat "Smoking status"
label define smkstat 1 "Current smoker" 2 "Ex-smoker" 3 "Never smoker"
label value smkstat smkstat

gen exercise = EXERCISELONG
replace exercise =. if EXERCISELONG == 8
label var exercise "Length of exercise"
label define exercise 0 "None" 1 "Less than 15 min" 2 "15-30 min" 3 "31-60 min" 4 "More than
60 min"
label value exercise exercise

gen routinedist = ROUTINEDISTGRP
replace routinedist = . if ROUTINEDISTGRP ==8
label var routinedist "Distance to rountine care in Quartiles"
label define routinedist 1 "0-2 Km (=< Q1)" 2 ">2-20 Km (Q1-Q2)" 3 ">20-51 Km (Q2-Q3)" 4 "
>51 Km ( >Q3)"
label value routinedist routinedist

///Work place exposures///
**Exposure to molds**
```

```
gen molds = 1 if i_MOLDS == 1
replace molds = 0 if i_MOLDS == 2
replace molds =. if i_MOLDS==888
label var molds "Have you ever been exposed to mold at work?"
label define molds 1 "Yes" 0 "No"
label value molds molds

**Mold exposure frequency***
gen mold_freq = 1 if i_MOLDS == 2
replace mold_freq = 2 if i_MOLDHOWOFTEN == 4
replace mold_freq = 3 if i_MOLDHOWOFTEN == 2|i_MOLDHOWOFTEN ==3
replace mold_freq = 4 if i_MOLDHOWOFTEN == 1
label var mold_freq "Frequency of exposure - Molds"
label define mold_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label value mold_freq mold_freq

**Exposure to diesel fumes
gen dieselfumes = i_DIESELFUMES
replace dieselfumes=. if i_DIESELFUMES ==888
label var dieselfumes "Have you ever been exposed to diesel fumes at work?"
label define dieselfumes 1 "Yes" 2 "No"
label value dieselfumes dieselfumes

***Diesel exposure frequency***
gen diesel_freq = 1 if i_DIESELFUMES == 2
replace diesel_freq = 2 if i_DIESELHOWOFTEN == 4
replace diesel_freq = 3 if i_DIESELHOWOFTEN == 2|i_DIESELHOWOFTEN == 3
replace diesel_freq = 4 if i_DIESELHOWOFTEN == 1
label var diesel_freq "Frequency of diesel - Diesel fumes"
label define diesel_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label value diesel_freq diesel_freq

**Exposure to smoke from stubble burning**
gen smokestubble = i_SMOKESTUBBLE
replace smokestubble =. if i_SMOKESTUBBLE == 888
label var smokestubble "Have you ever been exposed to smoking from stubble burning from
work?"
label define smokestubble 1 "Yes" 2 "No"
label value smokestubble smokestubble

***Smokestub exposure frequency***
gen smoke_freq = 1 if i_SMOKESTUBBLE == 2
replace smoke_freq = 2 if i_SMOKEHOWOFTEN == 4
replace smoke_freq = 3 if i_SMOKEHOWOFTEN == 3| i_SMOKEHOWOFTEN == 2
replace smoke_freq = 4 if i_SMOKEHOWOFTEN == 1
label var smoke_freq "Frequency of exposure - Smokestubble"
```

```
label define smoke_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label value smoke_freq smoke_freq

**Exposure to welding fumes**
gen welding = i_WELDING
replace welding =. if i_WELDING ==888
label var welding "Have you ever been exposed to welding fumes at work?"
label define welding 1 "Yes" 2 "No"
label value welding welding

***Welding fume exposure frequency***
gen weld_freq = 1 if i_WELDING == 2
replace weld_freq = 2 if i_WELDINGHOWOFTEN == 4
replace weld_freq = 3 if i_WELDINGHOWOFTEN == 3|i_WELDINGHOWOFTEN == 2
replace weld_freq = 4 if i_WELDINGHOWOFTEN == 1
label var weld_freq "Frequency of exposure - Welding fume"
label define weld_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label value weld_freq weld_freq

**Exposure to solvent fumes**
gen solvent = i_SOLVENT
replace solvent =. if i_SOLVENT ==888
label var solvent "Have you ever been exposed to solvent fumes at work?"
label define solvent 1 "Yes" 2 "No"
label value solvent solvent

***Solvent fumes exposure frequency***
gen solvent_freq = 1 if i_SOLVENT == 2
replace solvent_freq = 2 if i_SOLVENTHOWOFTEN == 4
replace solvent_freq = 3 if i_SOLVENTHOWOFTEN == 3| i_SOLVENTHOWOFTEN == 2
replace solvent_freq = 4 if i_SOLVENTHOWOFTEN == 1
label var solvent_freq "Frequency of exposure - solvent fumes"
label define solvent_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label val solvent_freq solvent_freq

**Exposure to radiation**
gen radiation = i_RADIATION
replace radiation =. if i_RADIATION ==888
label var radiation "Have you ever been exposed to radiation at work?"
label define radiation 1 "Yes" 2 "No"
label value radiation radiation

***Radiation exposure frequency***
gen rad_freq = 1 if i_RADIATION == 2
replace rad_freq = 2 if i_RADIATIONHOWOFTEN == 4
replace rad_freq = 3 if i_RADIATIONHOWOFTEN == 3|i_RADIATIONHOWOFTEN == 2
```

```
replace rad_freq = 4 if i_RADIATIONHOWOFTEN == 1
label var rad_freq "Frequency of exposure - Radiation"
label define rad_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label val rad_freq rad_freq

**Exposure to wood dust**
gen wooddust = 1 if i_WOODDUST == 1
replace wooddust = 0 if i_WOODDUST == 2
replace wooddust =. if i_WOODDUST ==888
label var wooddust "Have you ever been exposed to wood dust at work?"
label define wooddust 1 "Yes" 0 "No"
label value wooddust wooddust

***Wooddust exposure frequency***
gen wood_freq = 1 if i_WOODDUST == 2
replace wood_freq = 2 if i_WOODHOWOFTEN == 4
replace wood_freq = 3 if i_WOODHOWOFTEN == 3|i_WOODHOWOFTEN == 2
replace wood_freq = 4 if i_WOODHOWOFTEN == 1
label var wood_freq "Frequency of exposure - Wood dust"
label define wood_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label val wood_freq wood_freq

**Exposure to grain dust**
gen graindust = i_GRAINDUST
replace graindust =. if i_GRAINDUST ==888
label var graindust "Have you ever been exposed to grain dust at work?"
label define graindust 1 "Yes" 2 "No"
label value graindust graindust

***Grain dust exposure frequency***
gen grain_freq = 1 if i_GRAINDUST == 2
replace grain_freq = 2 if i_GRAINHOWOFTEN == 4
replace grain_freq = 3 if i_GRAINHOWOFTEN == 3|i_GRAINHOWOFTEN == 2
replace grain_freq = 4 if i_GRAINHOWOFTEN == 1
label var grain_freq "Frequency of exposure - Grain dust"
label define grain_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label val grain_freq grain_freq

**Exposure to asbestos dust**
gen asbestosdust = 1 if i_ASBESTOSDUST == 1
replace asbestosdust = 0 if i_ASBESTOSDUST == 2
replace asbestosdust =. if i_ASBESTOSDUST ==888
label var asbestosdust "Have you ever been exposed to asbestos dust at work?"
label define asbestosdust 1 "Yes" 0 "No"
label value asbestosdust asbestosdust
```

```
***Asbestos dust exposure frequency***
gen asbes_freq = 1 if i_ASBESTOSDUST == 2
replace asbes_freq = 2 if i_ASBESHOWOFTEN == 4
replace asbes_freq = 3 if i_ASBESHOWOFTEN == 3|i_ASBESHOWOFTEN == 2
replace asbes_freq = 4 if i_ASBESHOWOFTEN == 1
label var asbes_freq "Frequency of exposure - Asbestos"
label define asbes_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label val asbes_freq asbes_freq

**Exposure to mine dust**
gen minedust = 1 if i_MINEDUSTEXP == 1
replace minedust = 0 if i_MINEDUSTEXP == 2
replace minedust =. if i_MINEDUSTEXP ==888
label var minedust "Have you ever been exposed to mine dust at work?"
label define minedust 1 "Yes" 0 "No"
label value minedust minedust

***Mine dust exposure frequency***
gen mine_freq = 1 if i_MINEDUSTEXP == 2
replace mine_freq = 2 if i_MINEHOWOFTEN == 4
replace mine_freq = 3 if i_MINEHOWOFTEN == 3|i_MINEHOWOFTEN == 2
replace mine_freq = 4 if i_MINEHOWOFTEN == 1
label var mine_freq "Frequency of exposure - Grain dust"
label define mine_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label val mine_freq mine_freq

****Exposure to oil/gas well fumes***
gen oilgas = i_OILGAS
replace oilgas =.  if i_OILGAS ==888
label var oilgas "Have you ever been exposed to oil/gas well fumes at work?"
label define oilgas 1 "Yes" 2 "No"
label value oilgas oilgas

***Oil & gas exposure frequency***
gen oilgas_freq = 1 if i_OILGAS == 2
replace oilgas_freq = 2 if i_OILGASHOWOFTEN == 4
replace oilgas_freq = 3 if i_OILGASHOWOFTEN == 3|i_OILGASHOWOFTEN == 2
replace oilgas_freq = 4 if i_OILGASHOWOFTEN == 1
label var oilgas_freq "Frequency of exposure - Oil & Gas"
label define oilgas_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label val oilgas_freq oilgas_freq

***Pesticide***
gen pest=0 if i_FUNGICIDE == 2 |i_HERBICIDE == 2 |i_INSECT == 2
replace pest = 1 if i_FUNGICIDE == 1 |i_HERBICIDE == 1 |i_INSECT == 1
label var pest "Exposure to pesticides"
```

```
label define pest 0 "No" 1 "Yes"
label val pest pest

gen fung = 1 if i_FUNGICIDE == 2
replace fung = 2 if i_FUNGHOWOFTEN == 4
replace fung = 3 if i_FUNGHOWOFTEN == 3|i_FUNGHOWOFTEN == 2
replace fung = 4 if i_FUNGHOWOFTEN == 1

gen herb = 1 if i_HERBICIDE == 2
replace herb = 2 if i_HERBHOWOFTEN == 4
replace herb = 3 if i_HERBHOWOFTEN == 3|i_HERBHOWOFTEN == 2
replace herb = 4 if i_HERBHOWOFTEN == 1

gen insect = 1 if i_INSECT == 2
replace insect = 2 if i_INSECTHOWOFTEN == 4
replace insect = 3 if i_INSECTHOWOFTEN == 3|i_INSECTHOWOFTEN == 2
replace insect = 4 if i_INSECTHOWOFTEN == 1

gen pest_freq = 1 if insect == 1 | herb ==1 | insect == 1
replace pest_freq = 2 if insect == 2 | herb == 2 | insect == 2
replace pest_freq = 3 if insect == 3 | herb == 3 | insect == 3
replace pest_freq = 4 if insect == 4 | herb == 4 | insect == 4
label var pest_freq "Frequency of exposure - pesticide"
label define pest_freq 1 "None" 2 "Occasionally" 3 "weekly - monthly" 4 "Daily"
label val pest_freq pest_freq

**allergic to house dust***
gen allehousedust = 1 if i_ALLERGYHOUSEDUST == 1
replace allehousedust = 0 if i_ALLERGYHOUSEDUST == 2
label var allehousedust "Have you ever had allergic reaction to house dust?"
label define allehousedust 1 "Yes" 0 "No"
label value allehousedust allehousedust

**allergic to cats***
gen allecat = 1 if i_ALLERGYCAT == 1
replace allecat = 0 if i_ALLERGYCAT == 2
label var allecat "Have you ever had allergic reaction to cat?"
label define allecat 1 "Yes" 0 "No"
label value allecat allecat

**allergic to dogs***
gen  alledog = 1 if i_ALLERGYDOG == 1
replace alledog = 0 if i_ALLERGYDOG == 2
label var alledog "Have you ever had allergic reaction to dog?"
label define alledog 1 "Yes" 0 "No"
label value alledog alledog
```

```
**allergic to grass***
gen allegrass = 1 if i_ALLERGYGRASSES == 1
replace allegrass = 0 if i_ALLERGYGRASSES == 2
label var allegrass "Have you ever had allergic reaction to grasses?"
label define allegrass 1 "Yes" 0 "No"
label value allegrass allegrass

**allergic to pollens***
gen allepollen = 1 if i_ALLERGYPOLLENS == 1
replace allepollen = 0 if i_ALLERGYPOLLENS == 2
label var allepollen "Have you ever had allergic reaction to pollens?"
label define allepollen 1 "Yes" 0 "No"
label value allepollen allepollen

**allergic to molds***
gen allemolds = 1 if i_ALLERGYMOLDS == 1
replace allemolds = 0 if i_ALLERGYMOLDS == 2
label var allemolds "Have you ever had allergic reaction to molds?"
label define allemolds 1 "Yes" 0 "No"
label value allemolds allemolds

***other allergies***
gen otheralle = 1 if i_OTHERALLERGY == 1
replace otheralle = 0 if i_OTHERALLERGY == 2
label var otheralle "Do you have any other allergies?"
label define otheralle 1 "Yes" 0 "No"
label value otheralle otheralle

***any allergy***
gen anyalle = 1 if ANYALLERGY == 1
replace anyalle = 0 if ANYALLERGY == 2
label var anyalle "Do you have any allergy to respiratory outcome?"
label define anyalle 1 "Yes" 0 "No"
label value anyalle anyalle

**age continues var**
gen age = i_AGE
replace age = . if i_AGE == 888
label var age "Age continues?"
label value age age

***Age new grp***
gen age_new = 1 if age<=64
replace age_new = 2 if age >=65
replace age =. if age==888
label var age_new "Age two categories"
```

```
label define age_new 1 "<= 64" 2 ">=65"
label value age_new age_new

**diabetes**
gen diabetes = 1 if i_DIABETES == 1
replace diabetes = 0 if i_DIABETES == 2
label var diabetes "Have you ever had allergic reaction to cat?"
label define diabetes 1 "Yes" 0 "No"
label value diabetes diabetes

**ethic background**
gen ethnicity = 1 if i_ETHNICBG == 1
replace ethnicity = 2 if i_ETHNICBG == 2
replace ethnicity = 3 if i_ETHNICBG == 3
replace ethnicity = 4 if i_ETHNICBG == 4
label var ethnicity "Ethnic background"
label define ethnicity 1 "Causian" 2 "First Nation" 3 "Metis" 4 "Other"
label value ethnicity ethnicity

**Coal usage**
gen coal = 1 if h_COAL == 1
replace coal = 0 if h_COAL == 2
label var coal "Do you use coal in house"
label define coal 1 "Yes" 0 "No"
label value coal coal

save new_data, replace

///Correlation between SES indicators///

use new_data, replace
spearman incomeadq edugrp inc_new extramoney edu3gp

********** Converting data back to wide format **********
reshape wide i_SEX BMIGRP  FARM  INCOMEADQ INCOMEADQR INCOMEGRP i_AGE
AGEGRP EDUGRP MARITALGRP i_EDUCATION h_HOUSEINCOME ///
i_FARM1STYR SMKSTAT i_ALCOHOL  i_HEARTDIS  EXERCISELONG
EVERASTHMADR i_EMPHYSEVER  i_CHRBRONEVER i_COPDEVER i_HEARTATT ///
i_SMOKESTUBBLE h_COAL i_DIESELFUMES i_WELDING i_SOLVENT i_OILGAS
i_HERBICIDES i_FUNGICIDES i_INSECT i_MOLDS hc_index ///
h_EXTRAMONEY i_ALLERGYHOUSEDUST i_ALLERGYCAT i_ALLERGYDOG
i_ALLERGYGRASSES i_ALLERGYPOLLENS i_ALLERGYMOLDS ///
i_OTHERALLERGY ANYALLERGY i_DIABETES h_PEOPLE h_BEDROOMS h_NATGAS
h_DAMP h_CIGARETTES h_SMELL i_RADIATION inc_new ///
i_WOODDUST i_GRAINDUST i_ASBESTOSDUST i_MINEDUSTEXP sex bmi farm income
income_ru incomeadq age age_new pest ///
```

```
agegrp edugrp maritalgrp farm1styr smkstat alcohol heartdis exercise smokestubble dieselfumes
welding solvent edu3gp ///
oilgas molds extramoney hhsize bedrooms natgas damp h_cigarette smell radiation wooddust
graindust asbestosdust ///
minedust CRC copd allehousedust allecat alledog allegrass allepollen allemolds c_index
mold_freq diesel_freq smoke_freq ///
weld_freq solvent_freq rad_freq wood_freq grain_freq asbes_freq mine_freq oilgas_freq
otheralle anyalle coal diabetes ///
heartatt asthma fung herb insect pest_freq ACOS, i(HOUSEID PERSONID) j(time)

save wide_data, replace

*** Numbers and Percentages****
 **For baseline**
 foreach var of varlist edugrp1 edu3gp1 extramoney1 incomeadq1 income1 inc_new1 quad
farm1 miz hhsize1 bedrooms1  ///
 molds1 natgas1 damp1 smell1 h_cigarette1 coal1 smoke_freq1 diesel_freq1 weld_freq1
solvent_freq1 rad_freq1 wood_freq1  ///
 grain_freq1 asbes_freq1 mine_freq1 oilgas_freq1 pest_freq1 smkstat1 alcohol1 exercise1
farm1styr1 momsmoke hc_index1 ///
 allecat1 alledog1 allegrass1 allepollen1 allemolds1 otheralle1 anyalle1 age_new1 agegrp1 ///
 sex1 maritalgrp1 bmi1 oneparentlung heartdis1 diabetes1 heartatt1 routinedist farmdustsmk1
chemfumes1 chemdust1{
 tab `var', m
 }


***Setting data for survival data****

gen time_frame = ReceiveDt2 - h_DATE
drop if CRC1==1 | CRC1==.
keep if BOTHSTUDIES ==1
stset time_frame, id(PERSONID) failure(CRC2==1)

gen inc = 1 if incomeadq1 >= 1
replace inc = 2 if incomeadq1 ==.
tab inc CRC2, col chi2

gen par_lung = 1 if oneparentlung == 0|oneparentlung == 1
replace par_lung = 2 if oneparentlung ==.
tab par_lung CRC2, col chi2

gen extmoney = 1 if extramoney1 >=1
replace extmoney =2 if extramoney1 ==.
tab extmoney CRC2, col chi2
```

```
gen h_attack = 1 if heartatt1 <=1
replace h_attack = 2 if heartatt1 ==.
tab h_attack CRC2, chi2

 ***Incidence by characteristics***
 foreach var of varlist edugrp1 edu3gp1 extramoney1 incomeadq1 income1 inc_new1 quad miz
farm1 hhsize1 bedrooms1 ///
 molds1 natgas1 damp1 smell1 h_cigarette1 coal1 smokestubble1 dieselfumes1 welding1
solvent1 radiation1 hc_index1 ///
 wooddust1 graindust1 asbestosdust1 minedust1 oilgas1 pest1 smkstat1 ///
 alcohol1 exercise1 farm1styr1 momsmoke anyalle1 allecat1 alledog1 allegrass1 allepollen1
agegrp1 age_new1 ///
 sex1 maritalgrp1 bmi1 oneparentlung heartdis1 diabetes1 heartatt1 routinedist {
 tab CRC2 `var' if CRC1 == 0 & BOTHSTUDIES==1, col m chi2
 }

 ***Incidence rate by categories****
 foreach var of varlist edugrp1 extramoney1 incomeadq1 income1 quad farm1 hhsize1
bedrooms1  molds1 natgas1 ///
 damp1 smell1 h_cigarette1 coal1 smokestubble1 dieselfumes1 welding1 solvent1 radiation1
wooddust1 graindust1 ///
 asbestosdust1 minedust1 oilgas1 pest1 smkstat1 alcohol1 exercise1 farm1styr1 momsmoke ///
 allecat1 alledog1 allegrass1 allepollen1 anyalle1 age1 agegrp1 ///
 sex1 maritalgrp1 bmi1 oneparentlung heartdis1 diabetes1 routinedist farmdustsmk1
chemfumes1 chemdust1{
 stptime, by(`var')
 }

/////Univariate Analysis/////

**** Unadjusted HR ****
foreach var of varlist edugrp1 extramoney1 incomeadq1 income1 income_ru1 quad farm1
hc_index hhsize1 bedrooms1  molds1 natgas1 ///
 damp1 smell1 h_cigarette1 coal1 smokestubble1 dieselfumes1 welding1 solvent1 radiation1
wooddust1 graindust1 ///
 asbestosdust1 minedust1 oilgas1 pest1 smkstat1 alcohol1 exercise1 farm1styr1 momsmoke ///
 allehousedust1 allecat1 alledog1 allegrass1 allepollen1 allemolds1 otheralle1 anyalle1 ethnicity
agegrp1 ///
 sex1 maritalgrp1 bmi1 oneparentlung heartdis1 diabetes1 routinedist farmdustsmk1
chemfumes1 chemdust1{
 stcox i.`var', vce(cluster HOUSEID)
 }

 stcox i_AGE1, vce(cluster HOUSEID)
 stcox i.agegrp1, vce(cluster HOUSEID)
 stcox i.age_new1, vce(cluster HOUSEID)
```

```
stcox ib1.sex1, vce(cluster HOUSEID)
stcox i.maritalgrp1, vce(cluster HOUSEID)
stcox ib2.edugrp1, vce(cluster HOUSEID) /*needed*/
stcox ib3.edu3gp1, vce(cluster HOUSEID) /*needed*/
stcox i.extramoney1, vce(cluster HOUSEID) /*needed*/
stcox ib4.incomeadq1, vce(cluster HOUSEID) /*needed*/
stcox ib4.inc_new1, vce(cluster HOUSEID) /*needed*/
stcox i.quad, vce(cluster HOUSEID) /*needed*/
stcox i.miz, vce(cluster HOUSEID)
stcox ib2.farm1, vce(cluster HOUSEID) /*needed*/
stcox i.hc_index1, vce(cluster HOUSEID) /*needed*/
stcox ib2.hhsize1, vce(cluster HOUSEID) /*needed*/
stcox ib2.bedrooms1, vce(cluster HOUSEID)/*needed*/
stcox ib2.molds1, vce(cluster HOUSEID) /*needed*/
stcox ib1.natgas1, vce(cluster HOUSEID) /*needed*/
stcox ib1.damp1, vce(cluster HOUSEID) /*needed*/
stcox i.smell1, vce(cluster HOUSEID) /*needed*/
stcox i.h_cigarette1, vce(cluster HOUSEID) /*needed*/
stcox i.coal1, vce(cluster HOUSEID) /*NOT needed*/
stcox ib2.smokestubble1, vce(cluster HOUSEID)
stcox ib2.dieselfumes1, vce(cluster HOUSEID)
stcox i.welding1, vce(cluster HOUSEID)
stcox i.solvent1, vce(cluster HOUSEID)
stcox i.radiation1, vce(cluster HOUSEID)
stcox ib2.wooddust1, vce(cluster HOUSEID)
stcox ib2.graindust1, vce(cluster HOUSEID)
stcox ib2.asbestosdust1, vce(cluster HOUSEID)
stcox ib2.minedust1, vce(cluster HOUSEID)
stcox i.oilgas1, vce(cluster HOUSEID)
stcox i.pest1, vce(cluster HOUSEID)
stcox i.smoke_freq1, vce(cluster HOUSEID)
stcox i.diesel_freq1, vce(cluster HOUSEID)
stcox i.weld_freq1, vce(cluster HOUSEID)
stcox i.solvent_freq1, vce(cluster HOUSEID)
stcox i.rad_freq1, vce(cluster HOUSEID)
stcox i.wood_freq1, vce(cluster HOUSEID)
stcox i.grain_freq1, vce(cluster HOUSEID)
stcox i.asbes_freq1, vce(cluster HOUSEID)
stcox i.mine_freq1, vce(cluster HOUSEID)
stcox i.oilgas_freq1, vce(cluster HOUSEID)
stcox i.pest_freq1, vce(cluster HOUSEID)
stcox ib3.smkstat1, vce(cluster HOUSEID)
stcox ib1.alcohol1, vce(cluster HOUSEID)
stcox ib0.exercise1, vce(cluster HOUSEID)
stcox i.allehousedust1, vce(cluster HOUSEID)
stcox i.allecat1, vce(cluster HOUSEID)
```

```
stcox i.alledog1, vce(cluster HOUSEID)
stcox i.allegrass1, vce(cluster HOUSEID)
stcox i.allepollen1, vce(cluster HOUSEID)
stcox i.allemolds1, vce(cluster HOUSEID)
stcox i.otheralle1, vce(cluster HOUSEID)
stcox i.anyalle1, vce(cluster HOUSEID)
stcox i.farm1styr1, vce(cluster HOUSEID)
stcox i.momsmoke, vce(cluster HOUSEID)
stcox i.oneparentlung, vce(cluster HOUSEID)
stcox i.heartdis1, vce(cluster HOUSEID) /*needed*/
stcox i.diabetes1, vce(cluster HOUSEID) /*needed*/
stcox i.heartatt1, vce(cluster HOUSEID) /*needed*/
stcox i.bmi1, vce(cluster HOUSEID) /*needed*/
stcox i.routinedist, vce(cluster HOUSEID) /*needed*/


**********CRC**********
**************************************************
*** Statistical selection/ variables of p-value <0.20 and clinically important (age, sex and
maritalstat***

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 i.molds1
i.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1, vce(cluster HOUSEID)

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 i.age_new1#c.time_frame, vce(cluster HOUSEID)

**************************************************
**********Testing for interactions**********

***age***
stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 i.age_new1#ib3.edu3gp1, vce(cluster HOUSEID)  /* age & edu
*/ /*not sig*/
```

```
stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 i.age_new1#i.extramoney1, vce(cluster HOUSEID)  /*age &
extra*/ /*not sig*/

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 i.age_new1#ib4.incomeadq1, vce(cluster HOUSEID)  /* age &
inc */ /*not sig*/

 ***sex***
stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 ib1.sex1#ib3.edu3gp1, vce(cluster HOUSEID)  /* age & edu */
/*not sig*/


stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 ib1.sex1#i.extramoney1, vce(cluster HOUSEID)  /* age & extra
*/ /*not sig*/

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 ib1.sex1#ib4.incomeadq1, vce(cluster HOUSEID)  /* age & inc
*/ /*not sig*/

***Age and Sex***
stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 ib1.sex1#i.age_new1, vce(cluster HOUSEID)  /* age & sex */
/*not sig*/

**smoking**
```

```
stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 i.smkstat1#i.edu3gp1, vce(cluster HOUSEID)

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 i.smkstat1#i.extramoney1, vce(cluster HOUSEID)

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1 i.smkstat1#i.incomeadq1, vce(cluster HOUSEID)
 margin

 ***smoking as confounder***
stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1, vce(cluster HOUSEID) /*with smoke*/
 estimate store smoke

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 i.bmi1 i.farm1styr1 i.oneparentlung
///
i.anyalle1 i.heartdis1 i.diabetes1, vce(cluster HOUSEID) /*without smoke*/
 estimate store no_smoke

 esttab smoke no_smoke, replace ci cells(`"b(fmt(2) star) ci( par("(" "-" ")"))"') eform


//////////////////////////FINAL MODEL - Model without any interaction//////////////////////

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1, vce(cluster HOUSEID)
 est store stat_crc
```

```
/////MODEL DIAGNOSTICS//////////

 *****Testing for proportionality assumption********
stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1, tvc(i.age_new1 i.bmi1 ib3.smkstat1) vce(cluster HOUSEID)

 linktest

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1, vce(cluster HOUSEID)

 estat phtest

 **Graphical test**
 stphplot, by (incomeadq1) adjust(age_new1 sex1 quad miz molds1 natgas1 hc_index1 ///
h_cigarette1 wooddust1 asbestosdust1 minedust1 smkstat1 bmi1 farm1styr1 oneparentlung ///
anyalle1 heartdis1 diabetes1)

 ***Goodness of fit***
stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1, vce(cluster HOUSEID)
 predict cs, csnell
 stset cs, failure (CRC2==1)
 sts gen H = na
 line H cs cs, sort ytitle("") legend(cols(1))

 estat concordance

 ***power estimation***
 power cox , hratio(1.57) sd(0.4) onesided r2(0.1837) eventprob(0.023) n(3013)
 power cox , hratio(1.07) sd(0.4) onesided r2(0.1837) eventprob(0.023) n(3013)

 **Coefficient plot of significant factors**
coefplot stat_crc, xline(1) omitted baselevels headings(0.natgas1 = "{bf:Natural Gas}" ///
1.smkstat1 = "{bf:Smoking status}" 0.anyalle1 = "{bf:Any allergy}" 0.oneparentlung = ///
"{bf:Parent had lung disease}" 0.diabetes1 = ///
"{bf:Diabetes}") keep(*.natgas1 *.smkstat1 *.anyalle1 *.oneparentlung *.diabetes1) eform
```

```
**Coefficient plot of SES**
coefplot stat_crc, xline(1) omitted baselevels headings(1.edu3gp1 = "{bf:Education}" ///
1.extramoney1 = "{bf:Financial strain}" 1.incomeadq1 = "{bf:Income level}") ///
keep(*.edu3gp1 *.extramoney1 *.incomeadq1) eform


 **********COPD**********
 ***Statistical selection***
use wide_data, replace

gen time_frame = ReceiveDt2 - h_DATE
drop if copd1==1 | copd1==.
keep if BOTHSTUDIES ==1
stset time_frame, id(PERSONID) failure(copd2==1)

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1, vce(cluster HOUSEID)
 est store stat_copd

 ***power estimation***
 power cox , hratio(2.65) sd(0.4) onesided eventprob(0.061) n(3013)
 power cox , hratio(1.93) sd(0.4) onesided eventprob(0.061) n(3013)


 **********Asthma*********
 ***Statistical selection***
 use wide_data, replace

gen time_frame = ReceiveDt2 - h_DATE
drop if asthma1==1 | asthma1==.
keep if BOTHSTUDIES ==1
stset time_frame, id(PERSONID) failure(asthma2==1)

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1, vce(cluster HOUSEID)
 est store stat_asthma


 ***power estimation***
```

```
 power cox , hratio(1.57) sd(0.4) onesided eventprob(0.023) n(3013)
 power cox , hratio(1.07) sd(0.4) onesided r2(0.1837) eventprob(0.023) n(3013)

**********ACOS**********
 ******Stat selection******
 use wide_data, replace
 gen time_frame = ReceiveDt2 - h_DATE
drop if ACOS1 == 1 | ACOS1 ==.
keep if BOTHSTUDIES ==1
stset time_frame, id(PERSONID) failure(ACOS2==1)

stcox i.age_new1 ib1.sex1 i.quad i.miz ib3.edu3gp1 i.extramoney1 ib4.incomeadq1 ib1.molds1
ib1.natgas1 i.hc_index1 ///
i.h_cigarette1 i.wooddust1 ib1.asbestosdust1 ib1.minedust1 ib3.smkstat1 i.bmi1 i.farm1styr1
i.oneparentlung ///
i.anyalle1 i.heartdis1 i.diabetes1, vce(cluster HOUSEID)
 est store stat_acos

 esttab stat_crc stat_copd stat_asthma stat_acos using stat1.rtf, replace ci cells(`"b(fmt(2) star) &
ci( par("(" "-" ")"))"') eform
```