Measurement and effect of income on colorectal cancer survival and the diagnostic interval

Laura Ellen Davis

Department of Epidemiology, Biostatistics and Occupational Health

Faculty of Medicine

McGill University, Montreal, Canada

August 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy (Ph.D)

Table of Contents

Abstract

Colorectal cancer (CRC) patients experiencing low income have worse outcomes throughout the cancer continuum. Income inequalities in the diagnostic interval, the time from presentation of symptoms to the healthcare system to cancer diagnosis, may partially explain these outcomes. With increasing interest in income-related differences in cancer outcomes, accurate measurement of income is imperative, and misclassification of income can result in wrong conclusions about the presence of income inequalities. The overarching goal of this thesis was to estimate income inequalities in survival and the diagnostic interval for CRC patients in Canada and to advance knowledge regarding how the measurement of income at the individual and neighbourhood levels impacts those estimated inequalities.

The **first manuscript** determines misclassification between individual- and neighbourhood-level income and their association with survival among CRC patients diagnosed from 1992 to 2017 in the Canadian Census Health and Environment Cohorts. I found very poor agreement between individual and neighbourhood income, with only 17% of respondents assigned to the same quintile (weighted kappa=0.18). Individual income had a greater effect on relative and additive survival than neighbourhood income. The interaction between individual and neighbourhood income demonstrated that those in the lowest individual and neighbourhood income quintiles were the most at risk for poor survival. The poor agreement between these two measures fed directly into the **second manuscript,** where I used probabilistic bias analyses to adjust for exposure misclassification bias resulting from using neighbourhood income as a proxy for individual income when examining 5-year survival. The bias analysis resulted in similar relative risks (RR) for bias-adjusted neighbourhood income compared to true individual income. For example, the bias-adjusted RR for the lowest income quintile compared to the highest

income quintile was 1.57 (95% SI 1.41-1.73) compared to 1.68 (95% CI 1.58-1.79) for individual income. In situations where individual income is unavailable, these results can be applied to similar populations to estimate individual income effects more accurately on survival. Moreover, these methods can be adapted to other contexts to estimate the effects of income in different cohorts and diseases.

The **third manuscript** uses quantile regression to investigate the association between income and the diagnostic interval using administrative data within a cohort of colon cancer patients diagnosed in Ontario. Due to data constraints, I was unable to obtain individual income or present bias-adjusted income; therefore, I used neighbourhood income as a proxy for individual income. I defined the diagnostic interval by identifying and categorizing encounters occurring more frequently in the 0-3 months compared to 24-27 months before diagnosis and used statistical process control to define lookback periods for each encounter category. The first healthcare encounter was the earliest encounter, and the diagnostic interval was defined as the number of days from the first healthcare encounter to diagnosis. For patients with symptomatic pathways, low income was associated with longer diagnostic intervals. For example, the 90^{th} percentile diagnostic interval was 15 days (95% CI 6-23) longer in patients with the lowest income compared to the highest. These findings reveal income inequities during the diagnostic phase of colon cancer that may be underestimated using neighbourhood income instead of individual income.

This thesis demonstrates income inequities in survival and the diagnostic interval among individuals with CRC in Canada. The results from this thesis also stress the importance of measuring individual income to accurately estimate income inequalities in cancer outcomes and

provides a potential solution to estimate individual-level income effects when only neighbourhood income is available.

Résumé

Les patients atteints de cancer colorectal (CRC) avec un faible revenu ont des résultats plus mauvais tout au long du continuum du cancer. Les disparités de revenus dans l'intervalle diagnostique, le temps entre la présentation des symptômes au système de soins de santé et le diagnostic du cancer, pourraient en partie expliquer ces résultats. Face à l'intérêt croissant pour les différences liées au revenu dans les résultats du cancer, une mesure précise du revenu est impérative et une mauvaise classification du revenu peut conduire à des conclusions erronées sur la présence d'inégalités de revenus. L'objectif global de cette thèse était de combler les lacunes dans la compréhension de la mesure et de l'effet du revenu individuel et de quartier sur la survie, et de comprendre les inégalités de revenus dans l'intervalle diagnostique pour les patients atteints de CRC.

Le premier manuscrit détermine la mauvaise classification entre le revenu individuel et celui du quartier et leur association avec la survie globale parmi les patients atteints de CRC diagnostiqués de 1992 à 2017 dans les cohortes de santé et d'environnement du recensement canadien. Nous avons constaté un très faible accord entre le revenu individuel et le revenu du quartier, seulement 17% des répondants étant assignés au même quintile (kappa pondéré=0.18). Le revenu individuel avait un effet plus important sur la survie relative et additive que le revenu du quartier. L'interaction entre le revenu individuel et le revenu du quartier a montré que les plus à risque de mauvaise survie étaient ceux se trouvant dans les quintiles de revenu individuel et de quartier les plus bas.

Ces résultats alimentent directement le deuxième manuscrit où nous avons utilisé des analyses de biais probabilistes pour ajuster le biais de mauvaise classification d'exposition

résultant de l'utilisation du revenu du quartier par rapport au revenu individuel lors de l'examen de la survie à 5 ans. L'analyse des biais a abouti à des risques relatifs (RR) similaires pour le revenu ajusté pour le biais par rapport au revenu individuel.

Dans le troisième manuscrit, nous avons utilisé les méthodes présentées dans le manuscrit 3 et la régression quantile pour étudier l'association entre le revenu du quartier et l'intervalle diagnostique. Pour les patients présentant des voies symptomatiques, vivre dans des communautés à faible revenu du quartier était associé à des intervalles diagnostiques plus longs par rapport aux patients vivant dans des communautés au revenu de quartier le plus élevé.

Dans l'ensemble, ces résultats démontrent des inégalités de revenus à la fois individuelles et de quartier dans la survie et l'intervalle diagnostique chez les patients atteints de CRC. De plus, nous démontrons l'importance de mesurer le revenu individuel dans les études sur le cancer et nous proposons une solution potentielle aux situations où seul le revenu du quartier est disponible.

List of abbreviations

AJCC	American Joint Committee on Cancer
ARCC	Canadian Centre for Applied Research in Cancer Control
CIHR	Canadian Institute of Health Research
CI	Confidence Interval
CRC	Colorectal Cancer
CT	Computerized Tomography
ED	Emergency Department
FIT	Fecal Immunochemical Test
FRQS	Fonds de Recherche Quebec - Sante
gFOBT	Guiaic Fecal Occult Blood Test
MRI	Magnetic Resonance Imaging
OR	Odds Ratio
PCCF+	Postal Code Conversion File
PET	Positron Emission Tomography
RDC	Research Data Centre
RR	Risk Ratio
SI	Simulation Interval
TNM	Tumour, node, metastasis
UICC	International Union Against Cancer

List of Tables

List of Figures

Acknowledgements

I am so grateful to have had the opportunity, time, and support from so many awesome people to complete this thesis. First, I am incredibly thankful to my supervisor, Dr. Erin Strumpf, for her academic supervision, guidance, and mentorship. Your insight and perspective brought me down to earth or out of the weeds, whichever one I needed at the time. Thank you for pushing me to think out of the box. I am a better researcher because of your mentorship. Thank you to my co-supervisor Dr. Alyson Mahar. We have worked together for many years now, and I am so glad you could come along on yet another journey with me. Your mentorship, guidance, patience, last-minute talks, and shared laughs have helped get me to where I am today.

I am very lucky to have had the opportunity to work with Dr. Sunil Patel and Dr. Hailey Banack, whose thoughtful feedback greatly improved this work. Thank you also to Dr. Renzo Calderon-Anyosa. I am inspired by your endless patience and inquisitive nature.

To my PhD cohort, we were dealt a tough hand with the Covid pandemic but somehow managed to make the best of it. You are all so incredibly talented, and I hope that our paths continue to cross even as we go our separate ways. Thank you to my friends for bringing lightness to an otherwise very serious and independent time. I couldn't have done it without the coffee and drink dates, hiking, concerts, movies, and laughs. Thank you, especially to Jill. Your perspective on real life and the ridiculousness of grad school makes me laugh and has helped me to not get bogged down in the details.

Thanks to my family for their never-ending support. Especially to my mom, Mary, I made it this far because you instilled in me a sense of independence and that I could accomplish anything. Thank you to my weird dog Summer, who makes me laugh every day. Adopting her

got me through one of the toughest periods in my life. To Mike. Thank you for listening to some incredibly boring conversations with patience and (somehow) interest. You are truly the best, and I look forward to what life looks like together outside of the PhD.

Finally, this thesis is in loving memory of my Auntie Martha. You've always been immensely proud of my academic achievements, and I wish you could have seen it to the end.

Statement of financial support

I was very fortunate to receive several sources of financial support throughout my PhD. My central funding source was the Canadian Institute for Health Research (CIHR) Doctoral Award, which funded years 2 to 5 of my PhD. My first year was funded through the Studentship Award from the Canadian Centre for Applied Research in Cancer Control (ARCC) and the Doctoral Award from the Fonds de Recherche – Sante Quebec (FRQS). I also received supplementary support through research assistantships from my co-supervisor Alyson Mahar, funded through the Canadian Cancer Society new researcher award for the MEGAN-CAN project.

Contribution of authors

The individual manuscripts that make up my thesis are listed below. The data used in the analyses come from Statistics Canada linked datasets (manuscripts 1 and 2) and ICES (chapter 5 and manuscript 3). I developed the research questions and study design for all manuscripts with the support of my supervisors. I was solely responsible for merging the data, completing the analyses, and generating the draft versions of the manuscripts.

Dr. Erin Strumpf provided guidance from the protocol development stage onward. She helped me refine my research questions and build analysis plans for all manuscripts. Dr. Strumpf reviewed all manuscripts in detail and provided valuable feedback.

Dr. Alyson Mahar also provided guidance from the protocol development stage onward. She helped me refine the research questions and build analysis plans for all manuscripts. Dr. Mahar reviewed all manuscripts in detail and provided valuable feedback.

Dr. Sunil Patel contributed to Chapter 5 and Manuscript 3. He provided guidance on creating the diagnostic interval and helped interpret diagnosis and procedure codes and results. Dr. Patel reviewed the final chapter 5 and manuscript 3 in detail.

Dr. Hailey Banack provided technical expertise on the quantitative bias analysis in manuscript 2. She helped with the study design and methods. Dr. Banack reviewed manuscript 2 and provided valuable feedback.

Dr. Renzo Calderon-Anyosa checked the coding for the quantitative bias analysis and reviewed the final manuscript for manuscript 2.

Manuscript 1: Davis LE, Mahar AL, Strumpf EC. Agreement between individual and neighbourhood income measures in patients with colorectal cancer in Canada. JNCI: Journal of the National Cancer Institute. 2023 Jan 27:djad017.

Manuscript 2: Davis LE, Banack HR, Calderon-Anyosa R, Strumpf EC, Mahar AL. Validity and probabilistic bias analysis for individual income exposure misclassification by neighbourhood in colorectal cancer patients. Submitted to the International Journal of Epidemiology.

Manuscript 3: Davis LE, Strumpf EC, Patel S, Mahar AL. Income inequalities in time to colon cancer diagnosis. Submitted to Cancer Medicine.

Statement of originality

I attest that the work in this thesis is my own and represents an original and timely contribution to income inequities in colon and rectal cancer outcomes. My supervisors, committee members and co-authors provided indispensable guidance throughout the research process, but the overarching research questions and analyses reflected in this series of manuscripts are my own and were motivated by important gaps in the existing literature regarding income inequities and the diagnostic interval in colorectal cancer.

Chapter 1 : Introduction

The economic, social, and physical environments in which people live are shaped by the distribution of power, prestige, and resources. When these distributions are unequal, resulting in unfair, unacceptable and avoidable differences in health, we get health inequities.[1,2] The current literature on inequity in cancer care has established that differences in income influence cancer screening, stage at diagnosis and survival.[3–5] Measurement of income is particularly important since different measures can result in different findings and conclusions as to whether inequities are present.[6–8] Neighbourhood-level income measures are often used to approximate individual measures when individual data are either unavailable or restricted due to data confidentiality. However, individual and neighbourhood-level socioeconomic factors work through different pathways and may affect outcomes differently.[9–11] For example, individual income might affect cancer outcomes through material and social resources, such as income to pay for peripheral costs of cancer care (e.g. out of hospital prescriptions, parking and child care).[9,12] On the other hand, neighbourhood income might influence cancer outcomes through features of the physical environment, such as easier access to primary care or neighbourhood social support.[9,12] Some studies have examined differences between individual and neighbourhood-level income but few in a Canadian cancer context due to limited data availability and linkages.[10,13,14]

Colorectal cancer (CRC) is the third most commonly diagnosed cancer in men and women and the second leading cause of cancer death worldwide.[15,16] Understanding the role of income is especially important in CRC, where there are opportunities to decrease disparities through screening. In Canada, despite provincial screening programs, over 50% of patients are diagnosed at a late stage, and many are part of low-income groups, leading to further inequalities in stage at diagnosis and survival for those patients.[17–19] While inequalities in CRC screening are

well documented, less is understood of inequalities along the diagnostic pathway or the time from presentation of symptoms or screening to diagnosis. Shortening the diagnostic interval for patients experiencing low income could improve outcomes later in the cancer care continuum and improve equity in cancer outcomes.[13] A simplified conceptual model of the association between income, the diagnostic interval and survival is outlined in Figure 1.

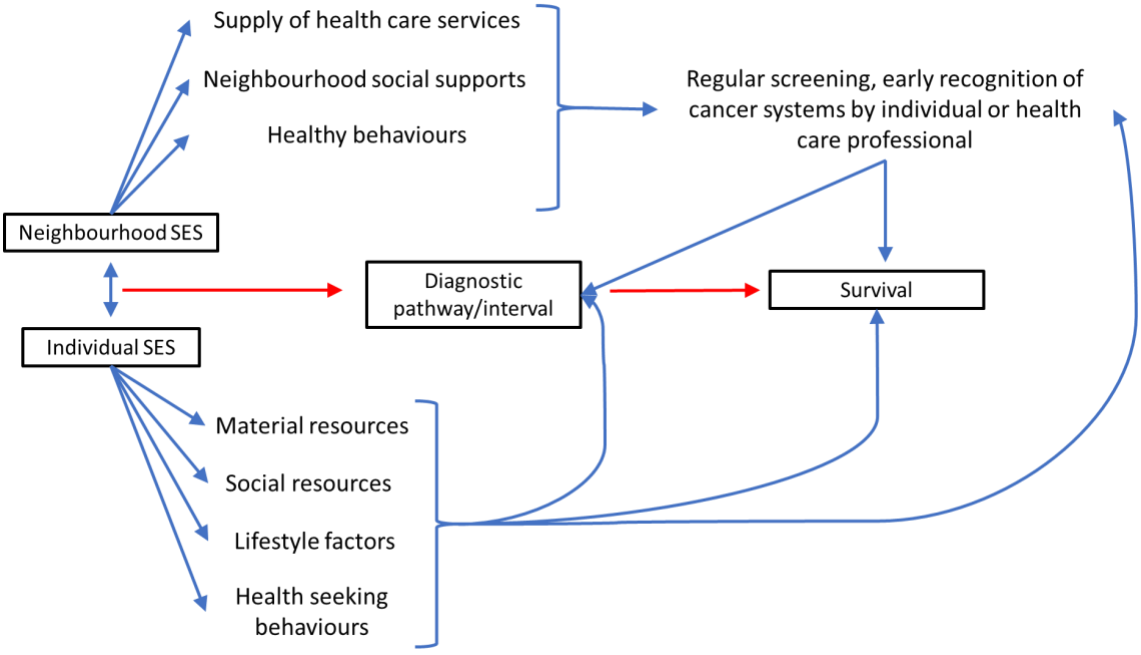


FIGURE 1.1. CONCEPTUAL MODEL FO THE ASSOCIATIONS BETWEEN SOCIOECONOMIC STATUS AND THE DIAGNOSTIC PATHWAYS/INTERVALS AND SURVIVAL

1.1 Research objectives

The overarching goal of this thesis was to fill several gaps in knowledge regarding income inequities that occur in colorectal cancer patients within the Canadian healthcare system. Specifically, to understand the measurement of income in studies using routinely collected administrative datasets and to evaluate and correct for the misclassification of individual income by neighbourhood income. I further assessed the effect of income on the diagnostic interval. The thesis consists of three objectives:

Objective 1. To estimate misclassification between income measured at the individual and neighbourhood level and to measure the association between individual and neighbourhood income independently and jointly on overall survival in a cohort of colorectal patients (Manuscript 1).

Objective 2. To treat neighbourhood income as misclassified individual income and adjust for exposure misclassification bias on the effect of 5-year mortality using quantitative bias analysis (Manuscript 2).

Objective 3. over (Manuscript 3).

1.2 Thesis structure

This manuscript-based thesis contains seven chapters. In **Chapter 1**, I introduce the overall rationale and research objectives. In **Chapter 2**, I provide an overview of the current literature on inequalities in cancer care, specifically income inequalities. I also provide contextual information on colorectal cancer epidemiology, the diagnostic pathways and intervals, and how income is usually measured in health studies in Canada. In **Chapter 3**, I present manuscript 1 (objective 1), where I performed a cohort study of Canadian colorectal cancer patients, using Statistics Canada linked datasets, to describe the misclassification between individual and neighbourhood income and estimate the effect of income on survival. Using the same data from Manuscript 1, I present Manuscript 2 (objective 2) in **Chapter 4**, where I apply quantitative bias analysis to demonstrate the adjusted effects of income on 5-year survival resulting from using neighbourhood income as a proxy for individual income. In **Chapter 5**, using administrative data from ICES (formerly the Institute of Clinical Evaluative Sciences), I describe the methodology I used to create the diagnostic interval in a cohort of colon cancer patients that was then used in the following chapter. In **Chapter 6,** I present Manuscript 3

(objective 3), where I evaluate the association between income and the diagnostic interval. Finally, **Chapter 7** summarises the results of the three objectives and offers implications, strengths, limitations, and suggestions for future research based on my collective findings.

Chapter 2 : Literature review

This chapter provides an overview of the current literature regarding inequities in cancer care, focusing specifically on income inequalities. I will provide contextual information on colorectal cancer epidemiology, diagnostic pathways, and the diagnostic interval, threading in the effect of income throughout. I will also give an overview of income measurement in health studies.

2.1 Inequity in cancer care

The terms inequity and inequality are often used interchangeably but have different definitions that can reflect different ways of thinking.[20] Health inequities are unfair, unacceptable and avoidable health differences resulting from the unequal distribution of power, prestige and resources across groups.[1,2] Stigma, bias and structural racism all contribute to health inequity. On the other hand, health inequalities are measured differences in health or health care; they can result in inequity, but not all inequalities are inequitable. Throughout this thesis, I will use the term inequality when measured differences are discussed. In contrast, I will use inequity when complex concepts related to social structure, power and injustice are involved and when differences are interpreted. Since these concepts are often difficult to tease apart, I use inequity more often as this term captures the unfair differences inherent when inequalities are present in cancer care by income.

Socioeconomic factors play a powerful role in determining health outcomes, and frameworks describing the social determinants of health and their causes have been described extensively.[1,21,22] The World Health Organization's Framework on Social Determinants of Health shows how socioeconomic positions reflect people's place within social hierarchies and,

in turn, shape specific determinants of health status.[23] In cancer specifically, the social, economic, environmental, historical, and political contexts in which individuals we live shape our outcomes across the cancer care continuum, contributing to the risk of developing and diagnosing cancer, the likelihood of receiving effective and timely treatment and the probability of dying from cancer.[24–27] The current literature on inequality and inequity in cancer care has established unequal cancer outcomes based on income, education, sex, ethnicity and race, and geography.[24] Research on cancer outcomes as a function of other determinants, such as sexual orientation, gender identity and immigration status, is less robust.[24,28,29]

The importance of social determinants of health and their impact on cancer outcomes have been recognized by cancer programs worldwide. Cancer programs across Canada, England, and Australia have all implemented cancer plans with the primary goal of reducing social inequities in the cancer system.[30–33] For example, one of the main goals of the most recent cancer plan set by Ontario Health Cancer Care Ontario describes reducing barriers to cancer care and improving health equity so Ontarians have equal access to cancer care regardless of what resources they have.[31]

2.1.1 Income inequity in cancer care

Barriers to navigating the healthcare system from cancer diagnosis to receiving treatment to palliative care exist at the patient, provider and health system levels.[34,35] These barriers are prevalent for all patients, but they disproportionately affect individuals facing structural inequities, such as poverty, leading to inequalities in cancer outcomes.[36,37] Individuals experiencing low income have been shown to have poorer outcomes across the cancer care continuum. For example, individuals with low income are more likely to participate in health behaviours, such as smoking and inactivity, have increased incidence of CRC, lower screening

rates, later stage at diagnosis and poorer survival.[3,4,24] Worse outcomes for those experiencing low income can occur through many complex and intersecting pathways. For example, jobs that pay less are less likely to have paid time off for medical appointments or sick leave, and comprehensive health benefits that cover prescription drugs potentially resulting in less timely and appropriate cancer diagnosis and treatment.[38] Pathways to worse cancer outcomes can occur as early as childhood, with evidence suggesting that adults who experienced childhood poverty have an increased risk of mortality from cancer compared to their peers.[39]

In the following sections, I will discuss in more detail income inequities that occur in CRC outcomes, such as incidence, stage at diagnosis and mortality and what is known about income inequities during the diagnostic phase of CRC.

2.2 Colorectal cancer incidence, stage at diagnosis, survival, and the effect of income

All manuscripts in the thesis focus on cohorts of Canadian colorectal or colon cancer patients. I conducted this research in CRC patients because this cancer is treatable if caught early, and early detection is possible through population-based organized screening programs in most provinces and territories in Canada.[40] Moreover, the development of CRC is influenced by factors such as smoking, diet and physical activity, which can be more common among those experiencing low income; thus, rates of CRC tend to be higher among individuals with low income.[41,42] This gives a greater opportunity to improve equity in outcomes by improving early detection across groups. In this section, I describe CRC epidemiology and the effect of income on CRC outcomes along the cancer care continuum.

CRC is the third most commonly diagnosed cancer in men and women, accounting for 12.7% and 10.0% of all cancer cases in Canadian males and females, respectively.[15,16,43] In Canada in 2022, the 5-year prevalence for CRC was 76,820 cases.[43] CRC is also a significant

contributor to mortality, accounting for 11% of all cancer-induced deaths in Canada.[44] Low income is associated with increased age-adjusted CRC prevalence and incidence.[43,45,46] Increased CRC incidence is associated with risk factors such as poor diet, low physical activity, smoking and alcohol consumption, all of which are more prevalent in individuals experiencing low income.[41,42]

Cancer stage is a classification system that incorporates the size of the tumour, which parts of the organ have cancer and whether and where the cancer has spread. The international standard for measuring stage is the tumour, node, metastasis (TNM) staging system of the American Joint Committee on Cancer (AJCC) and the International Union Against Cancer (UICC).[47] Increasing stage represents increasing growth of the tumour. Stage at diagnosis is a well-established prognostic factor in CRC management.[48] For instance, the 5-year survival rate for localized CRC (Stage I and II) can be as high as 92%, whereas, for advanced CRC (Stage IV), survival rates are approximately 11%.[44,49] In the Canadian context, a considerable proportion of CRC cases are diagnosed at a late stage (stage III or IV), with 49% of patients diagnosed at advanced stages in 2018.[44] Individuals experiencing low income are more likely to be diagnosed with CRC at a later stage than their high-income counterparts.[50–52] Consequently, they also consistently experience reduced survival compared to their higher-income peers.[5,53,54] There are many aspects along the cancer care continuum that can influence survival, such as stage at diagnosis, receipt of appropriate and timely treatment, and access to clinical trials, all of which tend to be worse among individuals who experience low income.[48,55–58] These statistics underscore the need for strategies that address socioeconomic disparities and promote early detection of CRC.

2.3 Diagnosing colorectal cancer and the effect of income

Chapters 5 and 6 of this thesis focus on defining the diagnostic interval and describing inequalities in the diagnostic interval by income. The following two sections will describe screening, diagnostic tests, and diagnostic pathways to colorectal cancer diagnosis. I will focus on diagnostic practices in Ontario, as this is where my study takes place. I will also summarise the existing literature describing the diagnostic interval and the effect of income.

CRC in Canada is diagnosed through one of two pathways: symptomatic and asymptomatic or screen-detected. In Ontario, Ontario Health Cancer Care Ontario outlines recommended pathways for symptomatic and screen-detected cancers, and similar guidelines exist in other Canadian provinces, the UK and Australia.[59–61] Regardless of the pathway, patients should receive a colonoscopy, either after a positive screening test or when they present with symptoms. I will describe each pathway in the following two sections. A simplified graphic representing symptomatic and screen-detected pathways is provided in Figure 2.[59]

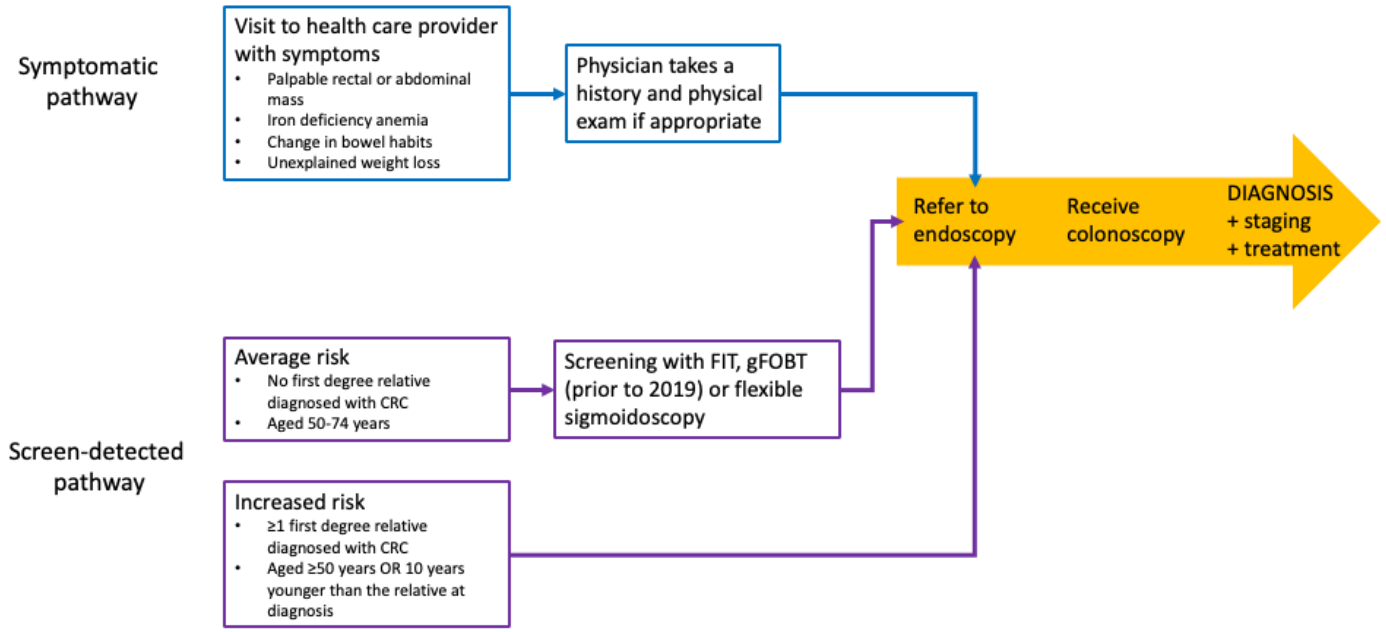


FIGURE 2.1. SIMPLIFIED COLORECTAL CANCER DIAGNOSTIC PATHWAYS (ADAPTED FROM ONTARIO HEALTH CANCER CARE ONTARIO COLORECTAL CANCER DIAGNOSIS PATHWAY MAP)

2.3.1 Screen-detected colorectal cancer

Screening for CRC is an important strategy for early detection that can catch the disease at an early stage, improve the effectiveness of treatment and increase survival rates.[62,63] Typically, screening is recommended for asymptomatic adults aged 50 to 75 or high-risk individuals with a family history of CRC or pre-existing conditions using one of the following tests: guaiac-based fecal occult blood test (gFOBT), fecal immunochemical test (FIT), flexible sigmoidoscopy, or colonoscopy.[60,64] FIT is the recommended screening test for most people and is a safe at-home test that checks the stool for tiny amounts of blood. The gFOBT test is similar to FIT but requires a restricted diet before the test. Since December 2019, Ontario has switched to FIT and no longer uses gFOBT kits. Flexible sigmoidoscopy uses a small, flexible tube with a camera to look inside the rectum and sigmoid colon. A colonoscopy is similar to a

sigmoidoscopy but can view the entire colon. Both flexible sigmoidoscopies and colonoscopies can take biopsies and remove polyps (abnormal growths).[64]

In Canada, all but one province (Quebec) and one territory (Nunavut) have implemented organized CRC screening.[65] Organized screening programs administer promotional, recruitment and reminder strategies to invite eligible individuals to be screened. Where organized screening programs are not available, screening services are provided opportunistically by a primary care provider. In April 2008, Ontario implemented the population-based ColonCancerCheck screening program organized through Ontario Health Cancer Care Ontario.[64] Low-risk individuals aged 50 to 74 without a history of CRC and high-risk individuals of any age that have a family history of CRC are invited to get screened.[64]

In Ontario, the pathway for screen-detected CRC begins with an invitation to screening or an opportunistic screening test at a primary care visit and is limited to eligible patients. Screen-eligible individuals receive either a FIT, flexible sigmoidoscopy or colonoscopy.[64] Screening options differ by risk. Low-risk individuals aged 50 to 74 without a family history of CRC are recommended FIT every two years. Moderate-risk individuals of any age with a family history of CRC before the age of 60 are recommended a colonoscopy every ten years, and high-risk individuals of any age with a family history of CRC that occurred before the age of 60 are recommended a colonoscopy every five years. Following an abnormal FIT or gFOBT result, individuals will require a colonoscopy. Following an abnormal colonoscopy, individuals should receive a diagnosis of CRC and proceed to staging and treatment. The Canadian Association of Gastroenterologists and Ontario Health Cancer Care Ontario recommend that colonoscopies be performed within eight weeks of an abnormal FIT.[66]

Inequalities in receipt of screening and screening pathways exist by income and other socioeconomic factors.[4,67] For example, in Canada, living in areas with low income is associated with having never been screened for CRC.[67] Moreover, while organized screening programs have been shown to increase screening practices in the general population, there is no evidence that the screening programs either increase or decrease income inequalities.[68,69]

2.3.2 Symptomatic colorectal cancer

Despite screening programs, most diagnosed patients initially present with noticeable signs and symptoms indicating the presence of CRC. The diagnostic process for patients with symptomatic CRC typically begins when a patient presents to a health care provider, most commonly a primary care provider.[59,70] Often, symptoms do not appear until the cancer has progressed to more advanced stages. Patients may present with a range of symptoms that can vary considerably based on the size and location of the cancer in the colon. Most commonly, patients present with changes in bowel habits, such as diarrhea or constipation, that last more than a few days.[71] Other symptoms might include rectal bleeding or blood in the stool, persistent abdominal discomfort such as cramps, gas pain or feeling that the bowel doesn't empty completely. Weight loss without a known reason, weakness or fatigue and nausea or vomiting can also indicate CRC. Importantly, these symptoms can also be caused by a number of other conditions, which is one of the significant challenges in diagnosing symptomatic CRC.[71,72]

Depending on the stage of the disease, symptomatic patients can present with mild or alarm symptoms. Up to 30% of CRC patients will have their first presentation in the Emergency Department (ED), usually with acute bowel obstruction symptoms or perforation requiring emergency surgery.[46,73] This type of presentation is associated with worse outcomes, such as later stage at diagnosis and worse survival.[74] Non-emergent patients might more commonly

present to a primary care provider, where they will be assessed through a detailed history and physical exam, which may include a digital rectal exam, abdominal examination and/or blood tests.[59,75] If the physician suspects CRC, they will refer them to an endoscopist or other appropriate specialist to receive a colonoscopy. If a colonoscopy is not possible or additional information is needed, imaging tests such as CT scans, MRI or PET scans may be used.[70] Following confirmation of CRC, the patient will begin staging and treatment.

2.4 The diagnostic interval and the effect of income

The diagnostic interval is the time from a patient's initial presentation to the healthcare system, either via a screening test or with symptoms suggestive of CRC, to the cancer diagnosis.[76–78] The diagnostic interval is an important determinant of patient outcomes. Longer diagnostic intervals can result in a later stage at diagnosis, further delays to treatment, worse overall survival and increased patient anxiety.[79] Monitoring and reducing the diagnostic interval is crucial for promoting earlier cancer diagnosis and improving outcomes. Some research has suggested that divergent diagnostic pathways and longer diagnostic intervals are more likely to occur in patients of lower socioeconomic positions, but little work has been done to explore this relationship.[80,81]

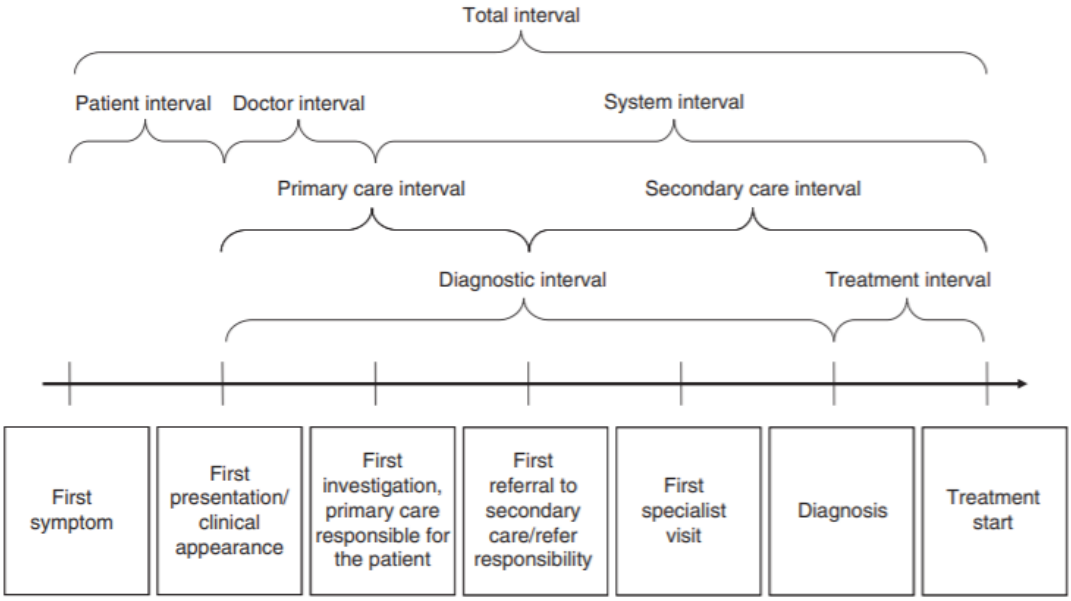


FIGURE 2.2. DESCRIPTION OF THE MILESTONES AND TIME INTERVALS ALONG THE DIAGNOSTIC ROUTE FROM THE FIRST SYMPTOM UNTIL THE START OF TREATMENT (FIGURE FROM WELLER ET AL., 2012)

Standardized time intervals across the diagnostic pathways have been defined to promote consistency across the literature (Figure 3).[82] The difficulty in defining the diagnostic interval for research using administrative health data lies in identifying the date of the first presentation to the healthcare system. Previous studies have used lookback periods of one or two years to look for CRC-related diagnostic procedures, such as colonoscopy, imaging, or consultations. This approach might overlook early misdiagnoses and result in artificially shortened intervals. Recently, Groome and colleagues developed a methodology that addresses some of these limitations, using administrative data to define the diagnostic interval.[76,83] This approach adopts variable lookback periods for each type of patient encounter based on statistical process control, offering a more nuanced, data-driven method to define diagnostic intervals.

2.5 Measurement of income

While income is a known social determinant of health, and experiencing low income has been shown in relation to worse cancer outcomes, difficulty still exists in accessing appropriate

measurements of income in routinely collected data, specifically for individual-level income measurements. The first two chapters of my thesis describe the misclassification between individual- and neighbourhood-level income and present a potential solution to obtain bias-adjusted income measures using quantitative bias analysis. The following two sections describe how income is usually measured in Canadian health studies and provide a brief overview of quantitative bias analysis.

2.5.1 Individual and area-level income

Income may be measured at the area, individual, or household levels. Area-level income is the mean or median aggregated over a census tract, dissemination area, or other geographic region. In Canadian health studies, area-level income is commonly used in the form of neighbourhood income obtained from the postal code conversion file plus (PCCF+).[84,85] The PCCF+ is created by Statistics Canada every five years along with the long-form census and assigns standard census geographic area-level variables, such as income and rural residence, to postal codes. Neighbourhood income from the PCCF+ represents the census-derived median income of all households in a census dissemination area, which includes approximately 400 to 700 people.[84,86] Conceptually, neighbourhood income is a place-based measure meant to represent the environment in which an individual lives and its related resources, such as community services, physical conditions, and access to health care.[87] In a cancer context, this might be understood as proximity to cancer centres or neighbourhood social supports.[12,88]

Individual income is measured for each person and represents an individual's ability to access material goods and services. Individual income is most commonly a measurement of their total income, through sources such as employment and other sources such as investments and governmental assistance. Individual income can be aggregated to the household level to obtain

the total spending abilities of a shared household. Household income is usually adjusted for household size using equivalence scales, commonly the square root of the household size.[89] Adjusting for household size accounts for the non-proportional growth in needs of the household. In Canada, individual and household income has only recently become available through Statistics Canada's Social Data Linkage Environment, with T1 tax files linked to the Canadian Census, the Canadian Community Health Survey, the Canadian Cancer Registry and other administrative data products.[90] Conceptually, individual income helps us understand structural inequalities and can be used to analyze income inequality and poverty at the individual level. In a Canadian cancer context, this might occur due to structural inequalities that benefit individuals who have more resources, such as the ability to pay for peripheral services, such as parking or childcare, or prescriptions not covered by the provincial health care plan.[12,88] Throughout this thesis, I will use individual and household income interchangeably.

2.5.2 Neighbourhood income as a proxy for individual income

While individual income has recently become available within Statistics Canada's Social Data Linkage Environment with Census and T1 tax files, it is still largely unavailable in many linked healthcare data sources across Canada. Moreover, accessing Statistics Canada's Social Data Linkage Environment provides its own unique challenges, and detailed information on physician billing has yet to be linked. Similar issues exist outside of Canada, with a lack of studies examining individual-level income inequalities in the US and the United Kingdom (UK).[91,92] For example, a systematic review examining socioeconomic status and cancer survival found that only 13 out of 74 studies examined individual or household income, and 10 of those studies were conducted in Denmark, Sweden or Norway.[54]

Due to the lack of availability, often stemming from issues related to privacy, many studies use neighbourhood income as a proxy for individual income, both in the cancer context and in other health services research.[14] Recent Canadian studies have shown that neighbourhood income is a poor proxy for individual income when examining outcomes in the general population, such as mortality, smoking and diabetes.[10,14] Studies in the US and the UK have also demonstrated poor agreement between individual and neighbourhood income.[91,92] Using neighbourhood income as a proxy for individual income often results in an underestimation of the effect of income on outcomes, which can lead to inaccurate conclusions about the presence of income inequalities.[10,14] Underestimation occurs when aggregation of individual measures are heterogenous within areas, decreasing the variation and resulting in an attenuation of effect.[93,94] This can potentially lead to researchers concluding that inequalities in the outcome do not exist by income if the proper interpretation of neighbourhood income is not considered.

2.5 Summary of knowledge gaps

Income inequities, measured through neighbourhood or individual income, impact colorectal cancer outcomes. The overarching goal of this thesis was to estimate income inequalities in survival and the diagnostic interval for CRC patients in Canada and to advance knowledge regarding how the measurement of income at the individual and neighbourhood levels impacts those estimated inequalities.

Objective 1 evaluates the agreement between individual and neighbourhood income quintiles among colorectal cancer patients diagnosed in Canada. While poor agreement between individual and neighbourhood income has been observed in the general Canadian population, studies have not yet examined this misclassification in a population diagnosed with cancer, where patients might be at a higher risk of experiencing low income. Moreover, I will compare

the same measure of income at the individual and neighbourhood levels and incorporate geography as a confounding characteristic. By addressing these gaps, I will provide a more nuanced understanding of the implications of using different income measures among colorectal cancer patients in Canada that can inform future studies.

Objective 2 treats neighbourhood income as misclassified individual income and uses quantitative bias analysis to estimate the bias-adjusted effect of income on 5-year mortality. This aim is motivated by the pervasive use of neighbourhood income as a proxy for individual income in cancer studies and the absence of individual income in routinely collected data sources. I present a potential solution for researchers to adjust estimated neighbourhood income effects in cancer studies as we wait for individual measures to become broadly available. Researchers can also modify this method for other disease sites and outcomes.

Objective 3 examines the association between neighbourhood income and the diagnostic interval. Unfortunately, due to data constraints, I was unable to obtain individual income or present bias-adjusted income; therefore, I used neighbourhood income as a proxy for individual income. Studies to date have produced inconclusive results regarding the association between income and the diagnostic interval, often failing to use a conceptual model and examining multiple variables in the same model without considering their role as confounders or causal pathway variables. In this study, my goal was to elucidate the impact of income on both the diagnostic pathways and interval. Due to restrictions on data access, I could not analyze the effect of individual income on the diagnostic interval or present bias-adjusted effects. This limitation underscores the pressing need for access to individual income data in large, linked datasets. Through this objective, I aim to advance our understanding of income disparities during cancer diagnosis, providing crucial insights into health inequities.

2.6 References

1. Marmot M, Friel S, Bell R, Houweling TA, Taylor S. Closing the gap in a generation: health equity through action on the social determinants of health. The Lancet. 2008 Nov 8;372(9650):1661–9.

2. Nelson A. Unequal treatment: confronting racial and ethnic disparities in health care. J Natl Med Assoc. 2002 Aug;94(8):666–8.

3. Manser CN, Bauerfeind P. Impact of socioeconomic status on incidence, mortality, and survival of colorectal cancer patients: a systematic review. Gastrointest Endosc. 2014 Jul 1;80(1):42-60.e9.

4. Pruitt SL, Shim MJ, Mullen PD, Vernon SW, Amick BC. Association of Area Socioeconomic Status and Breast, Cervical, and Colorectal Cancer Screening: A Systematic Review. Cancer Epidemiol Prev Biomark. 2009 Oct 1;18(10):2579–99.

5. Sinding C, Warren R, Fitzpatrick-Lewis D, Sussman J. Research in cancer care disparities in countries with universal healthcare: mapping the field and its conceptual contours. Support Care Cancer. 2014 Nov;22:3101–20.

6. Harper S, Lynch J. Methods for Measuring Cancer Disparities: Using Data Relevant to Healthy People 2010 Cancer-Related Objectives. Bethesda, MD: National Cancer Institute; 2005. (NCI Cancer Surveillance Monograph Series, Number 6).

7. Harper S, King NB, Meersman SC, Reichman ME, Breen N, Lynch J. Implicit Value Judgments in the Measurement of Health Inequalities. Milbank Q. 2010;88(1):4–29.

8. Harper S, Lynch J, Meersman SC, Breen N, Davis WW, Reichman ME. An Overview of Methods for Monitoring Social Disparities in Cancer with an Example Using Trends in Lung Cancer Incidence by Area-Socioeconomic Position and Race-Ethnicity, 1992–2004. Am J Epidemiol. 2008 Apr 15;167(8):889–99.

9. DeRouen MC, Schupp CW, Yang J, Koo J, Hertz A, Shariff-Marco S, et al. Impact of individual and neighborhood factors on socioeconomic disparities in localized and advanced prostate cancer risk. Cancer Causes Control. 2018;29(10):951–66.

10. Pichora E, Polsky JY, Catley C, Perumal N, Jin J, Allin S. Comparing individual and area-based income measures: impact on analysis of inequality in smoking, obesity, and diabetes rates in Canadians 2003–2013. Can J Public Health Rev Can Santé Publique. 2018 May 7;109(3):410–8.

11. Shariff-Marco S, Yang J, John EM, Sangaramoorthy M, Hertz A, Koo J, et al. Impact of Neighborhood and Individual Socioeconomic Status on Survival after Breast Cancer Varies by Race/Ethnicity: The Neighborhood and Breast Cancer Study. Cancer Epidemiol Prev Biomark. 2014 May 1;23(5):793–811.

12. DeRouen MC, Schupp CW, Koo J, Yang J, Hertz A, Shariff-Marco S, et al. Impact of individual and neighborhood factors on disparities in prostate cancer survival. Cancer Epidemiol. 2018;53:1–11.

13. Neal RD, Tharmanathan P, France B, Din NU, Cotton S, Fallon-Ferguson J, et al. Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes? Systematic review. Br J Cancer. 2015 Mar;112(1):S92–107.

14. Buajitti E, Chiodo S, Rosella LC. Agreement between area- and individual-level income measures in a population-based cohort: Implications for population health research. SSM - Popul Health. 2020 Apr 1;10:100553.

15. Morgan E, Arnold M, Gini A, Lorenzoni V, Cabasag CJ, Laversanne M, et al. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. Gut [Internet]. 2022 Sep 7 [cited 2022 Nov 7]; Available from: https://gut.bmj.com/content/early/2022/09/07/gutjnl-2022-327736

16. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022;72(1):7–33.

17. Canadian Cancer Statistics: A 2020 Special Report on Lung Cancer [Internet]. Toronto, ON: Canadian Cancer Society; 2020. Available from: cancer.ca/Canadian-Cancer-Statistics-2020-EN

18. Gupta S, Sussman DA, Doubeni CA, Anderson DS, Day L, Deshpande AR, et al. Challenges and Possible Solutions to Colorectal Cancer Screening for the Underserved. JNCI J Natl Cancer Inst [Internet]. 2014 Apr 1 [cited 2021 Nov 4];106(4). Available from: https://doi.org/10.1093/jnci/dju032

19. Carethers JM, Doubeni CA. Causes of Socioeconomic Disparities in Colorectal Cancer and Intervention Framework and Strategies. Gastroenterology. 2020 Jan 1;158(2):354–67.

20. Braveman P. What Are Health Disparities and Health Equity? We Need to Be Clear. Public Health Rep. 2014;129(Suppl 2):5–8.

21. Krieger N. Defining and investigating social disparities in cancer: critical issues. Cancer Causes Control. 2005 Feb 1;16(1):5–14.

22. Braveman P, Gottlieb L. The Social Determinants of Health: It's Time to Consider the Causes of the Causes. Public Health Rep. 2014;129(Suppl 2):19–31.

23. 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). [Internet]. Geneva: World Health Organization; 2010. Available from: https://apps.who.int/iris/bitstream/handle/10665/44489/9789241500852_eng.pdf?sequence=1&isAllowed=y

24. Sinding C, Warren R, Fitzpatrick-Lewis D, Sussman J. Research in cancer care disparities in countries with universal healthcare: mapping the field and its conceptual contours. Support Care Cancer. 2014 Nov;22:3101–20.

25. Ward E, Jemal A, Cokkinides V, Singh GK, Cardinez C, Ghafoor A, et al. Cancer disparities by race/ethnicity and socioeconomic status. CA Cancer J Clin. 2004 Apr;54:78–93.

26. Gilbert SM, Pow-Sang JM, Xiao H. Geographical Factors Associated With Health Disparities in Prostate Cancer. Cancer Control J Moffitt Cancer Cent. 2016 Oct;23(4):401–8.

27. Nolen LB, Braveman P, Dachs JNW, Delgado I, Gakidou E, Moser K, et al. Strengthening health information systems to address health equity challenges. Bull World Health Organ. 2005 Aug;83(8):597–603.

28. Quinn GP, Sanchez JA, Sutton SK, Vadaparampil ST, Nguyen GT, Green BL, et al. Cancer and lesbian, gay, bisexual, transgender/transsexual, and queer/questioning (LGBTQ) populations. CA Cancer J Clin. 2015 Oct;65:384–400.

29. Davis LE, Bogner E, Coburn NG, Hanna TP, Kurdyak P, Groome PA, et al. Stage at diagnosis and survival in patients with cancer and a pre-existing mental illness: a meta-analysis. J Epidemiol Community Health. 2020 Jan 1;74(1):84–94.

30. Cancer Australia. Australian Cancer Plan [Internet]. Surry Hills, NSW: Cancer Australia; 2022. Available from: file:///Users/lauradavis/Downloads/ACP-Public-Consultation---Draft-ACP-Summary-Form_1666138457.pdf

31. Cancer Care Ontario. Ontario Cancer Plan 2019-2023. Toronto: Cancer Care Ontario; 2018.

32. CancerCare Manitoba. Delivering Excellence: 2016-2021 Manitoba Cancer Plan. Winnipeg: CancerCare Manitoba; 2015.

33. The NHS. The NHS long term plan [Internet]. London, UK; 2019 Jan. Available from: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf

34. Ahmed S, Shahid RK. Disparity in cancer care: a Canadian perspective. Curr Oncol. 2012 Dec;19(6):e376–82.

35. Lambert LK, Horrill TC, Beck SM, Bourgeois A, Browne AJ, Cheng S, et al. Health and healthcare equity within the Canadian cancer care sector: a rapid scoping review. Int J Equity Health. 2023 Jan 28;22(1):20.

36. Bourgeois A, Horrill TC, Mollison A, Lambert LK, Stajduhar KI. Barriers to cancer treatment and care for people experiencing structural vulnerability: a secondary analysis of ethnographic data. Int J Equity Health. 2023 Mar 30;22(1):58.

37. International Agency for Research on Cancer. Reducing social inequalities in cancer: evidence and priorities for research. Lyon, France: World Health Organization; 2019 Apr.

38. Dixit N, Rugo H, Burke NJ. Navigating a Path to Equity in Cancer Care: The Role of Patient Navigation. Am Soc Clin Oncol Educ Book. 2021 Jun;(41):3–10.

39. Raphael D. Poverty in childhood and adverse health outcomes in adulthood. Maturitas. 2011 May 1;69(1):22–6.

40. Keane MG, Johnson GJ. Early diagnosis improves survival in colorectal cancer. The Practitioner. 2012 Jul 1;256(1753):15–9.

41. Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: A quantitative overview of the epidemiological evidence. Int J Cancer. 2009;125(1):171–80.

42. Pampel FC, Krueger PM, Denney JT. Socioeconomic Disparities in Health Behaviors. Annu Rev Sociol. 2010 Aug;36:349–70.

43. Canadian Cancer Statistics Advisory in collaboration with the Canadian Cancer Society, S. Canadian Cancer Statistics: A 2022 special report on cancer prevalence [Internet]. Toronto, ON: Canadian Cancer Society; 2022 Nov [cited 2023 Jun 19]. Available from: https://cancer.ca/en/research/cancer-statistics/canadian-cancer-statistics

44. Canadian Cancer Statistics Advisory Committee. Canadian Cancer Statistics 2018 [Internet]. Toronto, ON: Canadian Cancer Society; 2018 [cited 2018 Oct 2]. Available from: http://www.cancer.ca/~/media/cancer.ca/CW/cancer%20information/cancer%20101/Canadian%20cancer%20statistics/Canadian-Cancer-Statistics-2018-EN.pdf?la=en

45. Tope P, Morais S, El-Zein M, Franco E, Malagón T. Differences in site-specific cancer incidence by individual- and area-level income in Canada from 2006-2011 [Internet]. 2023 [cited 2023 Jun 19]. Available from: https://doi.org/10.1101/2023.01.13.23284277

46. Decker KM, Lambert P, Bravo J, Demers A, Singh H. Time Trends in Colorectal Cancer Incidence Rates by Income and Age at Diagnosis in Canada From 1992 to 2016. JAMA Netw Open. 2021 Jul 19;4(7):e2117556.

47. Compton CC, Greene FL. The Staging of Colorectal Cancer: 2004 and Beyond. CA Cancer J Clin. 2004;54(6):295–308.

48. McPhail S, Johnson S, Greenberg D, Peake M, Rous B. Stage at diagnosis and early mortality from cancer in England. Br J Cancer. 2015 Mar;112(1):S108–15.

49. American Cancer Society. Colorectal Cancer Survival Rates [Internet]. American Cancer Society; [cited 2023 Jun 19]. Available from: https://www.cancer.org/cancer/types/colon-rectal-cancer/detection-diagnosis-staging/survival-rates.html

50. Booth CM, Li G, Zhang-Salomons J, Mackillop WJ. The impact of socioeconomic status on stage of cancer at diagnosis and survival: a population-based study in Ontario, Canada. Cancer. 2010 Sep 1;116:4160–7.

51. Clegg LX, Reichman ME, Miller BA, Hankey BF, Singh GK, Lin YD, et al. Impact of socioeconomic status on cancer incidence and stage at diagnosis: selected findings from the surveillance, epidemiology, and end results: National Longitudinal Mortality Study. Cancer Causes Control. 2009;20:417–35.

52. Coughlin SS. Social determinants of colorectal cancer risk, stage, and survival: a systematic review. Int J Colorectal Dis. 2020 Jun 1;35(6):985–95.

53. Boyd C, Zhang-Salomons JY, Groome PA, Mackillop WJ. Associations between community income and cancer survival in Ontario, Canada, and the United States. J Clin Oncol Off J Am Soc Clin Oncol. 1999 Jul;17(7):2244–55.

54. Afshar N, English DR, Milne RL. Factors Explaining Socio-Economic Inequalities in Cancer Survival: A Systematic Review. Cancer Control J Moffitt Cancer Cent. 2021 Apr 30;28:10732748211011956.

55. Afshar N, Dashti SG, te Marvelde L, Blakely T, Haydon A, White VM, et al. Factors Explaining Socio-Economic Inequalities in Survival from Colon Cancer: A Causal Mediation Analysis. Cancer Epidemiol Biomarkers Prev. 2021 Oct 1;30(10):1807–15.

56. Davis LE, Coburn NG, Hallet J, Earle CC, Liu Y, Myrehaug S, et al. Material deprivation and access to cancer care in a universal health care system. Cancer. 2020;126(20):4545–52.

57. Chow CJ, Habermann EB, Abraham A, Zhu Y, Vickers SM, Rothenberger DA, et al. Does Enrollment in Cancer Trials Improve Survival? J Am Coll Surg. 2013 Apr;216(4):774.

58. Hanna TP, King WD, Thibodeau S, Jalink M, Paulin GA, Harvey-Jones E, et al. Mortality due to cancer treatment delay: systematic review and meta-analysis. BMJ. 2020 Nov 4;371:m4087.

59. Ontario Health Cancer Care Ontario. Colorectal Cancer Diagnosis Pathway Map [Internet]. Toronto; 2020 Jan. Available from: https://www.cancercareontario.ca/sites/ccocancercare/files/assets/CRCDiagnosisPathwayMap_2020.pdf

60. Public Health England. Bowel cancer screening care pathway [Internet]. England: Public Health England; [cited 2023 Jun 30]. Available from: https://www.gov.uk/government/publications/bowel-cancer-screening-care-pathway/bowel-cancer-screening-care-pathway

61. Cancer Council Australia. Clinical practice guidelines for the prevention, early detection and management of colorectal cancer [Internet]. Cancer Council Australia; 2017 Oct [cited 2023 Jun 30]. Available from: https://wiki.cancer.org.au/australia/Guidelines:Colorectal_cancer

62. Wiegering A, Ackermann S, Riegel J, Dietz UA, Götze O, Germer CT, et al. Improved survival of patients with colon cancer detected by screening colonoscopy. Int J Colorectal Dis. 2016 May 1;31(5):1039–45.

63. Brenner H, Jansen L, Ulrich A, Chang-Claude J, Hoffmeister M. Survival of patients with symptom- and screening-detected colorectal cancer. Oncotarget. 2016 May 17;7(28):44695–704.

64. Cancer Care Ontario. ColonCancerCheck Screening Recommendations. 2016.

65. Colorectal Screening in Canada: 2021/2022 Environmental Scan [Internet]. Canadian Partnership Against Cancer; 2022. Available from: https://www.partnershipagainstcancer.ca/topics/colorectal-cancer-screening-in-canada-2021-2022/summary/

66. Paterson WG, Depew WT, Paré P, Petrunia D, Switzer C, van Zanten SJV, et al. Canadian consensus on medically acceptable wait times for digestive health care. Can J Gastroenterol. 2006 Jun;20(6):411–23.

67. Blair A, Gauvin L, Ouédraogo S, Datta GD. Area-Level Income Disparities in Colorectal Screening in Canada: Evidence to Inform Future Surveillance. Curr Oncol. 2019 Apr;26(2):128–37.

68. Blair A, Gauvin L, Strumpf EC, Datta GD. Impact of organized colorectal cancer screening programs on screening uptake and screening inequalities: A study of systematic- and patient-reliant programs in Canada. J Cancer Policy. 2020 Jun 1;24:100229.

69. Charters TJ, Strumpf EC, Sewitch MJ. Effectiveness of an organized colorectal cancer screening program on increasing adherence in asymptomatic average-risk Canadians. BMC Health Serv Res. 2013 Oct 29;13(1):449.

70. Guan Z, Webber C, Flemming JA, Mavor ME, Whitehead M, Chen BE, et al. Real-world colorectal cancer diagnostic pathways in Ontario, Canada: A population-based study. Eur J Cancer Care (Engl). 2022;31(5):e13603.

71. Vega P, Valentín F, Cubiella J. Colorectal cancer diagnosis: Pitfalls and opportunities. World J Gastrointest Oncol. 2015 Dec 15;7(12):422–33.

72. Astin M, Griffin T, Neal RD, Rose P, Hamilton W. The diagnostic value of symptoms for colorectal cancer in primary care: a systematic review. Br J Gen Pract. 2011 May 1;61(586):e231–43.

73. Sheringham JR, Georghiou T, Chitnis XA, Bardsley M. Comparing primary and secondary health-care use between diagnostic routes before a colorectal cancer diagnosis: Cohort study using linked data. Br J Cancer. 2014 Oct;111(8):1490–9.

74. Golder AM, McMillan DC, Horgan PG, Roxburgh CSD. Determinants of emergency presentation in patients with colorectal cancer: a systematic review and meta-analysis. Sci Rep. 2022 Mar 14;12(1):4366.

75. Hamilton W, Sharp D. Diagnosis of colorectal cancer in primary care: the evidence base for guidelines. Fam Pract. 2004 Feb 1;21(1):99–106.

76. Groome PA, Webber C, Whitehead M, Moineddin R, Grunfeld E, Eisen A, et al. Determining the Cancer Diagnostic Interval Using Administrative Health Care Data in a Breast Cancer Cohort. JCO Clin Cancer Inform. 2019 Dec 1;(3):1–10.

77. Hamilton L. The Diagnostic Interval of colorectal cancer patients in Ontario by degree of rurality. [Kingston, ON]: Queen's University; 2015.

78. Webber C. Availability and quality of colonoscopy resource and the colorectal cancer diagnostic interval [Internet] [PhD Thesis]. [Kingston, ON]: Queen's University; 2017. Available from: https://qspace.library.queensu.ca/handle/1974/6707

79. Ontario Health Cancer Care Ontario. Rectal Cancer Treatment Pathway Map [Internet]. Toronto; 2020 Jan. Available from:

https://www.cancercareontario.ca/sites/ccocancercare/files/assets/RectalCancerTreatmentPathwayMap_2020.pdf

80. Sikdar KC, Dickinson J, Winget M. Factors associated with mode of colorectal cancer detection and time to diagnosis: a population level study. BMC Health Serv Res. 2017 Dec;17(1):1–11.

81. Webber C, Jiang L, Grunfeld E, Groome PA. Identifying predictors of delayed diagnoses in symptomatic breast cancer: a scoping review. Eur J Cancer Care (Engl). 2017;26(2):e12483.

82. Weller D, Vedsted P, Rubin G, Walter FM, Emery J, Scott S, et al. The Aarhus statement: improving design and reporting of studies on early cancer diagnosis. Br J Cancer. 2012 Mar;106(7):1262–7.

83. Webber C, Flemming JA, Birtwhistle R, Rosenberg M, Groome PA. Colonoscopy resource availability and its association with the colorectal cancer diagnostic interval: A population-based cross-sectional study. Eur J Cancer Care (Engl). 2020;29(1):e13187.

84. R Wilkins. PCCF + Version 3G Users Guide: Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files. Cat. No. 82F0086-XDB. 2001.

85. Measuring Health Inequalities: A Toolkit: Area-Level Equity Stratifiers Using PCCF and PCCF+ [Internet]. Ottawa: Canadian Institutes of Health Information; 2018. Available from: https://www.cihi.ca/sites/default/files/document/cphi-toolkit-area-level-measurement-pccf-2018-en-web.pdf

86. Davis LE, Mahar AL, Strumpf EC. Agreement between individual and neighborhood income measures in patients with colorectal cancer in Canada. JNCI J Natl Cancer Inst. 2023 Jan 27;djad017.

87. Chokie M, Partridge MD. Low-Income Dynamics in Canadian Communities: A Place-Based Approach. Growth Change. 2008;39(2):313–40.

88. DeRouen MC, Schupp CW, Yang J, Koo J, Hertz A, Shariff-Marco S, et al. Impact of individual and neighborhood factors on socioeconomic disparities in localized and advanced prostate cancer risk. Cancer Causes Control. 2018;29(10):951–66.

89. What are equivalence scales? Paris: Organisation for Economic Co-operation and Development; 2011.

90. Trudeau R. Social Data Linkage Environment: IJPDS (2017) Issue 1, Vol 1:057, Proceedings of the IPDLN Conference (August 2016). Int J Popul Data Sci [Internet]. 2017 Apr 13 [cited 2023 Jun 29];1(1). Available from: https://ijpds.org/article/view/76

91. Moss JL, Johnson NJ, Yu M, Altekruse SF, Cronin KA. Comparisons of individual- and area-level socioeconomic status as proxies for individual-level measures: evidence from the Mortality Disparities in American Communities study. Popul Health Metr. 2021 Jan 7;19(1):1.

92. Ingleby FC, Belot A, Atherton I, Baker M, Elliss-Brookes L, Woods LM. Assessment of the concordance between individual-level and area-level measures of socio-economic deprivation in a cancer patient cohort in England and Wales. BMJ Open. 2020 Nov 26;10(11):e041714.

93. Abrahamowicz M, du Berger R, Krewski D, Burnett R, Bartlett G, Tamblyn RM, et al. Bias due to Aggregation of Individual Covariates in the Cox Regression Model. Am J Epidemiol. 2004 Oct 1;160(7):696–706.

94. Hanley GE, Morgan S. On the validity of area-based income measures to proxy household income. BMC Health Serv Res. 2008 Apr 10;8(1):79.

Chapter 3 : Manuscript 1. Agreement between individual and neighbourhood income measures in patients with colorectal cancer in Canada

3.1 Preface

There is a pervasive use of neighbourhood income as a proxy for individual income in health services research, including cancer studies. While some studies have demonstrated that neighbourhood income is a poor proxy for individual income in the general population, no studies have examined the agreement between individual and neighbourhood income in a cancer population, where individuals are at a higher risk of experiencing low income. In this manuscript, I compare the most commonly used neighbourhood income measure in Canada to individual household income in a cohort of colorectal cancer patients. To ensure comparability, I created the individual income measure to be as similar as possible to the neighbourhood income measure created by Statistics Canada. I additionally examined the effects of individual and neighbourhood income, independently and jointly, on survival. This manuscript will help cancer researchers better interpret income measurements and stresses the need for access to individual socioeconomic measures in administrative data sources.

This manuscript was peer-reviewed and published in the *Journal of the National Cancer Institute.*

Citation: Davis LE, Mahar AL, Strumpf EC. Agreement between individual and neighborhood income measures in patients with colorectal cancer in Canada. JNCI: Journal of the National Cancer Institute. 2023 May 1;115(5):514-22.

3.2 Manuscript 1

ABSTRACT

Introduction: With increasing interest in income-related differences in cancer outcomes, accurate measurement of income is imperative. Misclassification of income can result in wrong conclusions as to the presence of income inequalities. We determined misclassification between individual- and neighbourhood-level income and their association with overall survival among colorectal cancer (CRC) patients.

Methods. The Canadian Census Health and Environment Cohorts were used to identify CRC patients diagnosed from 1992 to 2017. We used neighbourhood income quintiles from Statistics Canada and created individual income quintiles from the same data sources to be as similar as possible. Agreement between individual and neighbourhood income quintiles was measured using cross-tabulations and weighted kappa statistics. Cox proportional hazards and Lin's semi-parametric hazards models were used to determine the effects of individual and neighbourhood income independently and jointly on survival. Analyses were also stratified by rural residence.

Results. 103,530 CRC patients were included in the cohort. There was poor agreement between individual and neighbourhood income with only 17% of respondents assigned to the same quintile (weighted kappa=0.18). Individual income had a greater effect on relative and additive survival than neighbourhood income when modeled separately. The interaction between individual and neighbourhood income demonstrated that the most at risk for poor survival were those in the lowest individual and neighbourhood income quintiles. Misclassification was more likely to occur for patients residing in rural areas.

Conclusion. Cancer researchers should avoid using neighbourhood income as a proxy for individual income, especially among patients with cancers with demonstrated inequalities by income.

Keywords. Income, socioeconomic factors, neoplasms, colorectal neoplasms, survival, health inequities, social determinants of health

INTRODUCTION

Colorectal cancer (CRC) is the third most commonly diagnosed cancer in both men and women and the second leading cause of cancer death worldwide.[1,2] For patients with CRC, income effects every step along the cancer continuum. Individuals experiencing low income are more likely to participate in health behaviours, such as smoking and inactivity, have increased incidence of CRC, lower screening rates, later stage at diagnosis and poorer survival.[3–5]

Accurate measurement of income is required to understand cancer inequalities and target interventions to reduce inequities. However, the choice of income measure is often determined by data availability instead of appropriateness. For example, individual-level income measurements are rarely available in population-based datasets therefore researchers often use area-based measures.[6] This is especially common in cancer studies, for example, a recent systematic review of socioeconomic status and cancer survival found that only 15 of 66 studies examining socioeconomic status (excluding education) measured individual socioeconomic status.[7] Using area-level income measurements as a proxy for individual-level income is potentially problematic for several reasons. If the area is large or heterogeneous, the association observed at the area-level will not reflect the association at the individual-level.[8] In addition, the mechanistic pathways through which individual and area-level socioeconomic factors effect outcomes may differ, resulting in different interpretations and interventions.[49–11] For example, individual income might affect cancer outcomes through material and social resources, such as the ability to pay for peripheral costs of cancer care like parking and child care.[9,12] On the other hand, neighbourhood income might influence cancer outcomes through features of the physical environment, such as proximity to cancer centres or neighbourhood social supports.[9,12]

Several studies have demonstrated poor agreement between individual and neighbourhood income in the public health literature, but none have examined this in cancer specifically, where patients might be at a higher risk of experiencing low income.[10,13–17] These studies often fail to compare the same measure of income at the individual and neighbourhood level resulting in inaccurate comparisons.[10,12,15] Moreover, few studies consider geography as a confounding characteristic. This is especially important since heterogenous geographical areas might increase the chance of ecologic fallacy.[10,13,15,18]

In this study we aimed to 1) estimate misclassification between income measured at the individual and neighbourhood level across rural and non-rural residences and 2) measure the association of individual and neighbourhood income independently and jointly on overall survival in a cohort of colorectal cancer patients.

METHODS

Data and Study Population

This was a retrospective cohort study of a subset of participants diagnosed with CRC from the 1991-2011 Canadian Census Health and Environment Cohorts (CanCHEC). The CanCHEC is a national population-based cohort derived by Statistics Canada that links census, vital statistics and cancer registry databases to follow individuals who respond to the long form census for mortality and cancer outcomes.[19] The Canadian Cancer Registry collects cancer data on all Canadians with cancer, except Quebec after 2010.[20] The census is performed every 5 years in Canada and approximately 1 in every 5 households complete the long form census. The current study used census data from 1991, 1996, 2001, 2006 and the National Household Survey (NHS) in 2011. The NHS was a voluntary survey performed in 2011 instead of the census.

Individuals are eligible to complete the long form census or NHS if they were residents of Canada on census day and were not living in an institution, such as nursing homes, penitentiaries, or group homes. The postal code conversation file (PCCF+) was linked to the CanCHEC using the participants' postal code at census to obtain neighbourhood income quintiles and rural residence.

We included Canadians over the age of 35 with a new diagnosis[6,24–26] of CRC (ICD-O-3 codes: C180, C182-C189, C199, C209) between January 1, 1992 and December 31, 2017. Individuals had to have completed at least one long form census in the ten years prior to diagnosis (census cycles 1991-2011). A ten-year time frame was chosen to maximize sample size while minimizing any changes in income or residence that may have occurred between the census and diagnosis. In cases where participants completed more than one census, the closest questionnaire to the cancer diagnosis was used. Individuals with a missing date of diagnosis, missing postal code at census, postal code that could not be linked to the PCCF+ and those with a death date before their diagnosis date were excluded.

Exposure

Income was measured at the individual and neighbourhood level. We created the individual level income measure to be as consistent as possible with neighbourhood level income which was obtained from the PCCF+ (summarized in Table 1).

TABLE 3.1 INDIVIDUAL AND NEIGHBOURHOOD INCOME DEFINITIONS

Income definition concepts	Individual income	Neighbourhood income
Data source	Census (self-reported in 1991, 1996 and 2001 and linked to tax records in 2006 and 2011)	PCCF+ (uses Statistics Census Profile Data)
Date of measurement	Date of census	Census year using postal code at time of census
Income	Total income for all household members	Total income in each EA/DA (calculated by multiplying the median income in each EA/DA by the number of households)
Household size adjustment	Statistics Canada single person equivalents	Sum of the single person equivalents of the EA/DA (calculated by multiplying the number of households by the single person equivalence separately for each household size of 1 to 5+ and summing)
Before/after tax	Before tax	Before tax
Quintiles	Created by authors in each weighted census population by CMA/CA/other region	Created by Statistics Canada in each weighted census population by CMA/CA/other region

Individual income was defined as adjusted before-tax household income using the long form census. All sources of income from the previous calendar year were summed for each household and adjusted for number of household members using the single person equivalence scales from Statistics Canada's low income cut offs.[21] Before-tax income was used for consistency across all census years as after-tax income was only available in 2011. Income was self-reported in years 1991, 1996 and 2001 with the option to consent to income tax linkage in 2006 and 2011. All individuals included in the cohort consented to the use of income tax linkage. Continuous individual income was adjusted to 2011 Canadian dollars using the Statistics Canada consumer price index.[22] **Individual income quintiles** were created by ordering individual household incomes within each census metropolitan area (CMA), census agglomeration (CA) and other region by province and census year within the full weighted census cohort, as opposed

to within the CRC cohort. We then divided each category into 5 equal groups to create quintiles specific for each CMA/CA/other region. CMAs are large urban areas of ≥100,000 people, CAs are smaller areas of ≥10,000 population and other regions incorporate urban fringe and rural areas.[23] CMA/CA specific quintiles takes into account differences in cost of living across regions.

Neighbourhood income quintiles were created by Statistics Canada for the PCCF+. The PCCF+ neighbourhood income is the most widely used area-level income measure in Canada.[24] The PCCF+ uses Statistics Canada Census Profile Data at the dissemination or enumeration area (DA/EA) level to calculate area-level adjusted household income.[25] EAs and DAs are Statistics Canada's smallest geographical area representing approximately 400-700 individuals per area.[26] Total income for each DA/EA is calculated by multiplying the median income of that area by the total number of households. This number is then adjusted for household size by dividing by the sum of the single-person equivalents of the DA/EA to obtain median household income per single person equivalent for each DA/EA. CMA/CA quintiles are then constructed by ranking DAs/EAs within each CMA/CA/other region by province from lowest to highest then dividing into fifths.[27] Statistics Canada did not create a neighbourhood income quintile variable in 2011 due to the use of the NHS instead of the census, therefore we assigned 2006 neighbourhood income quintiles to individuals who responded to the 2011 census.

Outcome

Death from any cause was defined according to the vital statistics database. Follow up time was defined as the number of days from the date of CRC diagnosis in the CCR to the date of death from any cause, end of study (December 31, 2019), or loss to follow-up. Loss to follow

up was defined as those without a death date that had at least four years of consecutively missing postal codes without a returning postal code.

Covariates

Individual characteristics obtained from the census were age at census and sex. Tumour location (colon or rectum) and stage at diagnosis were obtained from the CCR.[28] Stage was only presented for individuals diagnosed in 2010 or after, as this was when the CCR started prioritizing the routine collection of stage data for lung, colorectal, breast and prostate cancers.[29] Geographic characteristics were measured at the time of census and included province/territory of residence, whether the person had moved in the 5 years previous to the census, and residence in a rural area, defined according to the PCCF+ as residing in a census subdivision with a population of <1,000 and a population density of <400 persons per square kilometre.

Statistical analysis

We examined misclassification between individual and neighbourhood income quintiles using cross tabulations overall and stratified by rural residence. Weighted Kappa statistics were calculated to determine the degree of non-random concordance between the two income measures.[30] Continuous individual income was described within individual and neighbourhood income quintiles. The cohort was described using means, medians and interquartile ranges for continuous variables and proportions for categorical variables, overall, by census year and by individual and neighbourhood income quintiles. In keeping with data confidentiality guidelines from Statistics Canada, number of observations are rounded to the nearest 5.

Cox proportional hazards regression and Lin's semi-parametric additive risk models were used to determine the relative and additive associations between exposures and overall survival.

Four models were specified to estimate the association of individual income, neighbourhood income, individual and neighbourhood income, and the interaction between individual income and neighbourhood income with survival. The interaction models included terms for individual income, neighbourhood income and the interaction term between the two. Models were stratified by rural and non-rural residence. The reference for all models was the highest income quintile 5. Multivariable models were adjusted for age, sex, tumour location, census year and province/territory at census. We reported adjusted and unadjusted hazard ratios (HRs) or risk differences (RD) for additional deaths per 1,000 person years and their 95% confidence intervals (95% CI) and statistical significance was considered at $p>0.05$. Proportional hazards were evaluated for all models through graphical diagnostics of the weighted Schoenfeld residuals for the Cox models and the Kolmogorov-Smirnov and Cramer von Mises tests for the additive risk models. A sensitivity analysis was performed stratifying the cox proportional hazards models by stage at diagnosis to see if associations differed by stage.

Ethics approval was obtained from the McGill University Research Ethics Board and Statistics Canada. SAS (version 9.4) was used for all analyses except the additive risk regression models which were analysed using the timereg package in R.[31]

RESULTS

Study cohort

There were 122,040 adults aged 35 or older with a first CRC diagnosis between 1992 and 2017 and who responded to the long form census in the ten years before their cancer diagnosis. Exclusions included 14,145 with a missing postal code on the date of census, 4,345 with a postal code that could not be linked in the PCCF+ and therefore did not have information on

neighbourhood income, and 20 with a death date before their diagnosis date, resulting in a final cohort of 103,530. The average time from measurement of income to diagnosis was 4.9 years (standard deviation (SD) of 3 years) and was similar across individual and neighbourhood income quintiles. There were some differences in individual and tumour characteristics across census years (Supplementary Table 1).

Misclassification of individual and neighbourhood income

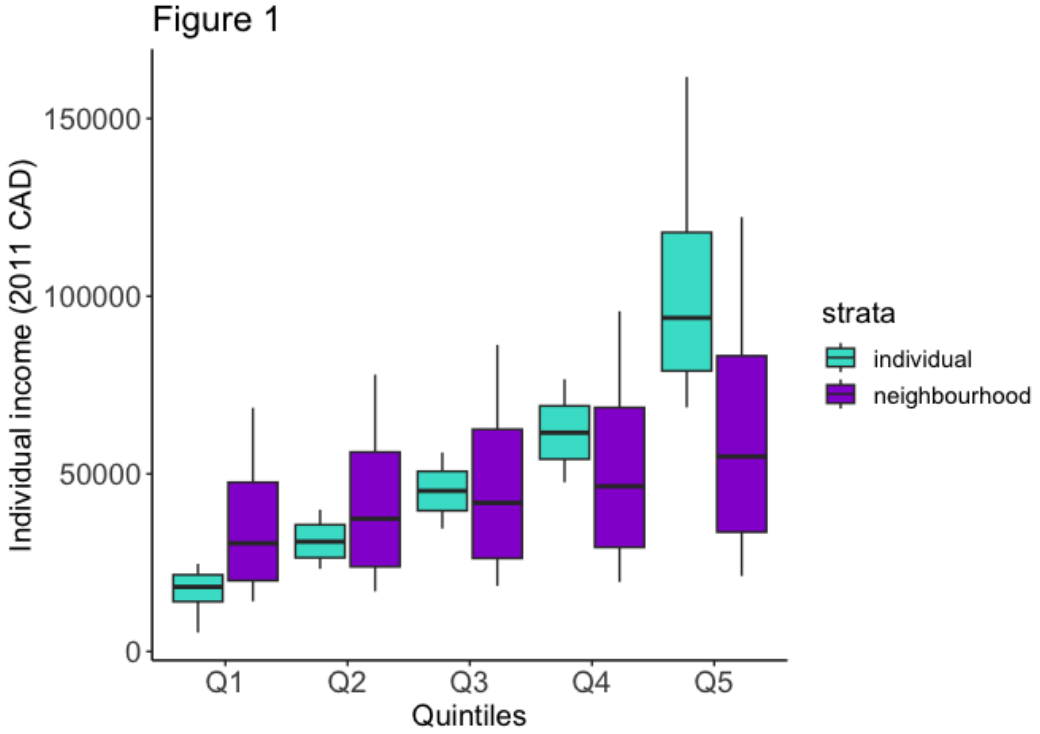
Overall agreement between individual and neighbourhood income quintiles at the time of census was poor (weighted kappa 0.18). Seventeen percent of patients were assigned to the same individual and neighbourhood income quintile and 54% were assigned either to the same quintile or one above or below. Individuals residing in rural areas were more likely to be misclassified than those in urban residences, with a weighted kappa of 0.11 compared to 0.20 in urban areas (Table 2, Supplementary Table 2). The range of individual income in 2011 Canadian dollars across individual income quintiles was wider and the variation within quintiles was smaller compared to neighbourhood income quintiles (Figure 1). Median individual income within individual income quintiles varied from $18,187 (IQR $14,051 – $21,591) in quintile 1 (Q1) to $93,902 (IQR $78,975 – $117,947) in quintile 5 (Q5), whereas neighbourhood quintiles varied from $30,451 (IQR $19,960 – $47,626) in Q1 to $54,882 (IQR $33,608 – $83,200) in Q5.

TABLE 3.2 AGREEMENT BETWEEN INDIVIDUAL AND NEIGHBOURHOOD INCOME QUINTILES (N, ROW PERCENT)

	Neighbourhood income quintile					
Individual income quintile	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Q1	7480 (32.3)	5380 (23.3)	4210 (18.2)	3430 (14.8)	2640 (11.4)	23140
Q2	5375 (22.9)	5520 (23.5)	4965 (21.1)	4255 (18.1)	3395 (14.4)	23505
Q3	3520 (17.8)	4335 (21.9)	4200 (21.2)	4170 (21.1)	3565 (18.0)	19790
Q4	2545 (13.8)	3540 (19.3)	3905 (21.3)	4215 (22.9)	4170 (22.7)	18375
Q5	1860 (9.9)	2635 (14.1)	3455 (18.5)	4330 (23.1)	6435 (34.4)	18715
Total	20780	21410	20735	20400	20205	103530

* Abbreviations: Q=quintiles.

FIGURE 3.1 INDIVIDUAL INCOME DISTRIBUTION BY INDIVIDUAL AND NEIGHBOURHOOD INCOME QUINTILES (WHISKERS REPRESENT 10TH AND 90TH PERCENTILES AS OPPOSED TO MINIMUM AND MAXIMUM VALUES DUE TO DATA CONFIDENTIALITY)



Characteristics by individual and neighbourhood income

A greater number of individuals with CRC were categorized in the lowest individual income quintile (Q1=23,145 vs Q5=18,715), while individuals were evenly spread across neighbourhood income quintiles (Q1=20,780 vs Q5=20,205)(Table 3). Patients in the lowest individual income quintile were older and more likely to be female compared to those in the highest income quintile (Median age 68 in Q1 vs 61 in Q5; Q1=54% female vs Q5=37%), whereas age and sex distributions were similar across neighbourhood income quintiles. Patients in Q1 for both individual and neighbourhood income were more likely to die during the follow up period compared to Q5, with a greater difference by individual income (Q1=71% vs Q5=51% for individual income and Q1=66% vs Q5=57% for neighbourhood income).

TABLE 3.3 INDIVIDUAL CHARACTERISTICS BY INDIVIDUAL AND NEIGHBOURHOOD INCOME QUINTILES (N, ROW PERCENT)

	Individual income quintiles					Neighbourhood income quintiles				
	Q1 (lowest quintile, N=23,145)	**Q2 (N=23,505)**	**Q3 (N=19,790)**	**Q4 (N=19,275)**	**Q5 (highest quintile, N=18,715)**	**Q1 (lowest quintile, N=20,780)**	**Q2 (N=21,410)**	**Q3 (N=20,735)**	**Q4 (N=20,400)**	**Q5 (highest quintile, N=20,205)**
Variable	**N (%)**	**N (%)**	**N (%)**	**N (%)**	**N (%)**	**N (%)**	**N (%)**	**N (%)**	**N (%)**	**N (%)**
Age at census (mean, SD)	68 (12)	67 (12)	64 (12)	62 (12)	61 (11)	65 (12)	65 (12)	65 (12)	64 (12)	64 (12)
Age at census										
35-44	1035 (4.7)	1095 (4.7)	1210 (6.1)	1165 (6.3)	1050 (5.6)	1065 (5.1)	1025 (4.8)	1140 (5.5)	1185 (5.8)	1140 (5.7)
45-54	2220 (9.6)	2175 (9.3)	2800 (14.1)	3585 (19.5)	4465 (23.9)	2730 (13.1)	2865 (13.4)	3105 (15.0)	3245 (15.9)	3295 (16.3)
55-64	4900 (21.2)	4410 (18.8)	4980 (25.2)	5400 (29.4)	6245 (33.4)	4945 (23.8)	5160 (24.1)	5205 (25.1)	5340 (26.2)	5280 (26.1)
65-74	7470 (32.3)	8840 (37.6)	6510 (32.9)	5100 (27.8)	4280 (22.9)	6630 (31.9)	6925 (32.4)	6355 (30.7)	6200 (30.4)	6085 (30.1)
75-84	5960 (25.7)	5810 (24.7)	3570 (18.0)	2580 (14.0)	2195 (11.7)	4350 (20.9)	4460 (20.8)	4050 (19.5)	3625 (17.8)	3630 (18.0)
85+	1560 (6.7)	1180 (5.0)	725 (3.7)	550 (3.0)	485 (2.6)	1065 (5.1)	975 (4.6)	880 (4.2)	810 (4.0)	770 (3.8)
Sex										
Male	10680 (41.2)	13230 (56.3)	11730 (59.3)	11040 (60.1)	11840 (63.3)	11255 (54.2)	11930 (55.7)	11740 (56.6)	11810 (57.9)	11785 (58.3)
Female	12460 (53.8)	10275 (43.7)	8060 (40.7)	7340 (39.9)	6875 (36.7)	9525 (45.8)	9475 (44.3)	9000 (43.4)	8590 (42.1)	8420 (41.7)
Province/Territory at census										
Newfoundland and Labrador	545 (2.4)	670 (2.8)	520 (2.6)	495 (2.7)	445 (2.4)	535 (2.6)	555 (2.6)	535 (2.6)	575 (2.8)	480 (2.4)
PEI	155 (0.7)	105 (0.5)	100 (0.5)	125 (0.7)	120 (0.6)	125 (0.6)	105 (0.5)	125 (0.6)	140 (0.7)	110 (0.6)
Nova scotia	960 (4.2)	990 (4.2)	850 (4.3)	745 (4.1)	810 (4.3)	900 (4.3)	855 (4.0)	880 (4.2)	860 (4.2)	865 (4.3)
New Brunswick	655 (2.8)	735 (3.1)	585 (3.0)	480 (2.6)	525 (2.8)	665 (3.2)	580 (2.7)	610 (3.0)	570 (2.8)	555 (2.8)
Quebec	5035 (21.8)	4810 (20.5)	3730 (18.8)	3320 (18.1)	3455 (18.5)	4525 (21.8)	4425 (20.7)	3995 (12.3)	3725 (18.3)	3675 (18.2)
Ontario	8935 (38.6)	9185 (39.1)	7905 (39.9)	7575 (41.2)	7625 (40.7)	7615 (36.7)	8375 (39.1)	8330 (40.2)	8390 (41.1)	8510 (42.1)
Manitoba	970 (4.2)	1010 (4.3)	940 (4.7)	845 (4.6)	880 (4.7)	890 (4.3)	1045 (4.9)	965 (4.7)	915 (4.5)	815 (4.0)
Saskatchewan	815 (3.5)	915 (3.9)	735 (3.7)	690 (3.8)	705 (3.8)	800 (3.8)	810 (3.8)	785 (3.8)	760 (3.7)	710 (3.5)
Alberta	2080 (9.0)	1850 (7.9)	1630 (8.2)	1465 (8.0)	1535 (8.2)	1745 (8.4)	1760 (8.2)	1760 (8.5)	1655 (8.1)	1640 (8.1)
British Columbia	2910 (12.6)	3160 (13.4)	2730 (13.8)	2555 (13.9)	2540 (13.6)	2910 (14.0)	2830 (13.2)	2645 (12.8)	2740 (13.4)	2780 (13.8)
Territories combined	75 (0.3)	80 (0.3)	70 (0.4)	75 (0.4)	80 (0.4)	70 (0.3)	70 (0.3)	105 (0.5)	75 (0.4)	65 (0.3)
Rural (pop<1000)										
Urban	17505 (75.6)	17560 (74.7)	15170 (76.7)	14270 (77.7)	14495 (77.4)	15935 (76.7)	16470 (76.9)	15610 (75.3)	15410 (75.5)	15580 (77.1)
Rural	5640 (24.4)	5945 (25.3)	4620 (23.3)	4105 (22.3)	4220 (22.6)	4845 (23.3)	4940 (23.1)	5130 (24.7)	4995 (24.5)	4625 (22.9)

Tumour location										
Rectal	7525 (32.5)	7665 (32.6)	6820 (34.4)	6275 (34.1)	6530 (34.9)	7005 (33.7)	7195 (33.6)	6960 (33.6)	6915 (33.9)	6740 (33.4)
Colon	15615 (67.5)	15840 (67.4)	12975 (65.6)	12105 (65.9)	12180 (65.1)	13775 (66.3)	14215 (66.4)	13780 (66.4)	13485 (66.1)	13465 (66.7)
Status										
Alive at end of follow up	6750 (29.2)	8035 (34.2)	8160 (41.21)	8485 (46.2)	9205 (59.2)	7090 (34.1)	8015 (37.4)	8240 (39.7)	8580 (42.1)	8705 (43.1)
Died	16395 (70.8)	15475 (65.8)	11635 (58.8)	9890 (53.8)	9510 (50.8)	13690 (65.9)	13395 (62.6)	12495 (60.3)	11820 (57.9)	11500 (56.9)
Stage at diagnosis*										
Stage 0-I	930 (18.8)	1285 (20.5)	1235 (21.9)	1190 (22.2)	1275 (23.8)	1030 (20.0)	1135 (20.5)	1185 (21.1)	1295 (22.7)	1270 (22.8)
Stage II	1170 (23.7)	1515 (24.2)	1320 (23.4)	1220 (22.8)	1130 (21.0)	1250 (24.2)	1300 (23.5)	1275 (22.7)	1310 (23.0)	1225 (22.0)
Stage III	1210 (24.5)	1555 (24.8)	1505 (26.7)	1450 (27.1)	1460 (27.3)	1315 (25.5)	1450 (26.2)	1475 (26.3)	1495 (26.2)	1450 (26.0)
Stage IV	905 (18.4)	1130 (18.0)	965 (17.1)	875 (16.3)	920 (14.2)	910 (17.7)	960 (17.4)	1000 (17.8)	945 (16.5)	980 (17.5)
Unknown	300 (6.0)	295 (4.7)	220 (3.9)	185 (3.5)	185 (3.4)	245 (4.8)	245 (4.4)	220 (3.9)	240 (4.2)	235 (4.2)
Missing	430 (8.7)	490 (7.8)	405 (7.2)	430 (8.1)	395 (7.4)	405 (7.9)	445 (8.0)	460 (8.2)	420 (7.4)	420 (7.6)

* Stage at diagnosis only for individuals diagnosed in 2010 or later

* Abbreviations: Q=quintile; pop=population.

Relative and additive survival by individual and neighbourhood income

Unadjusted individual income had a greater association with survival than neighbourhood income when modeled separately (Q1 vs Q5 individual HR 1.69 & RD 50.74, neighbourhood HR 1.20 & RD 19)(Figure 2, Table 4). After adjusting for individual covariates, the association of individual income on survival was attenuated while neighbourhood income remained similar (Q1 vs Q5 individual HR 1.36 & RD 32.23, neighbourhood HR 1.20 & RD 19). This could indicate that confounders are not appropriately controlled for at the neighbourhood level due to only controlling for individual-level covariates. To facilitate comparability between the two measures, unadjusted effects are reported moving forward. When both measures were included in the same model, the estimates for neighbourhood income were attenuated while individual income remained similar to it's unadjusted estimate. This suggests that some, but not all, of the effect of neighbourhood income on survival is accounted for by individual income.

FIGURE 3.2 5-YEAR KAPLAN MEIER SURVIVAL BY INDIVIDUAL AND NEIGHBOURHOOD INCOME QUINTILES

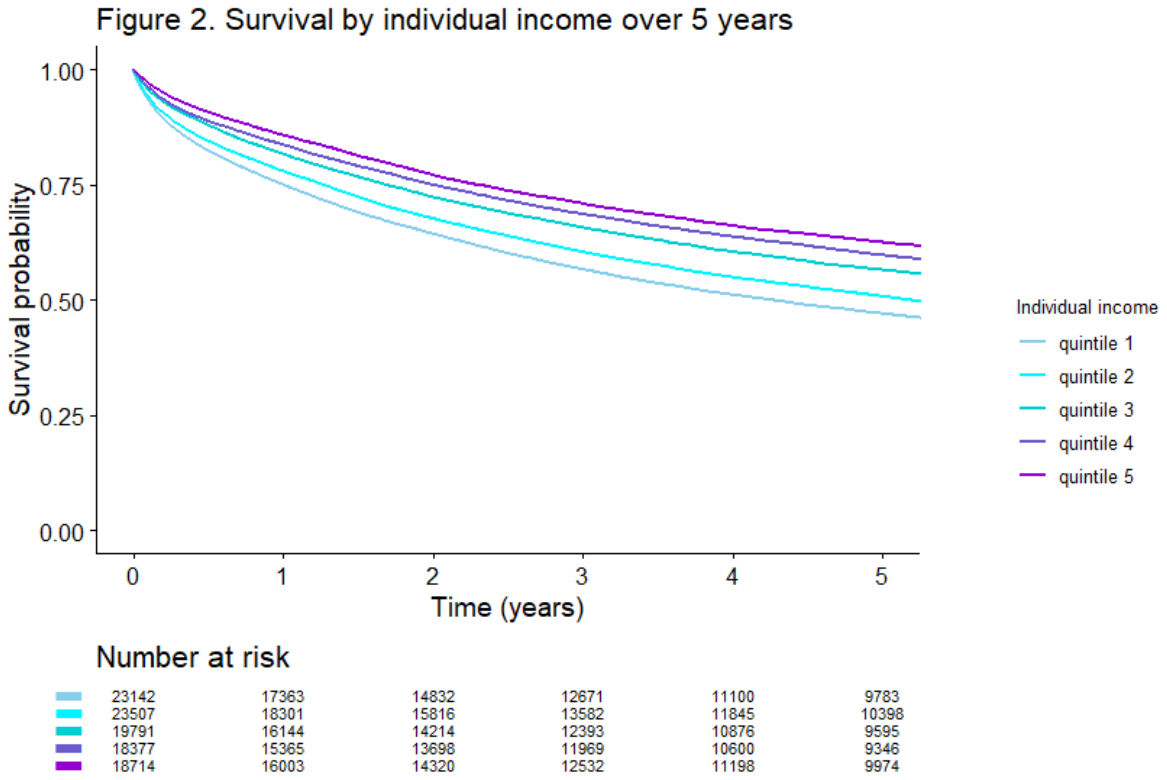


TABLE 3.4 COX PROPORTIONAL HAZARDS REGRESSION AND LIN'S SEMI-PARAMETRIC HAZARDS REGRESSION FOR THE ASSOCIATION BETWEEN INCOME AND SURVIVAL

	Cox regression model			**Lin's regression model**	
Model	**Unadjusted HR (95% CI)**	**Adjusted HR (95% CI)**		**Unadjusted RD (95% CI) per 1,000 py**	**Adjusted RD (95% CI) per 1,000 py**
Individual (Ref = Q5)					
Q1	1.69 (1.65-1.73)	1.36 (1.32-1.39)		50.74 (31.35-70.12)	32.23 (15.49-48.97)
Q2	1.52 (1.48-1.56)	1.23 (1.20-1.26)		37.60 (19.42-55.77)	19.86 (4.19-35.52)
Q3	1.26 (1.22-1.29)	1.12 (1.09-1.16)		17.99 (0.68-35.31)	9.45 (-5.36-24.26)
Q4	1.10 (1.07-1.13)	1.06 (1.03-1.09)		6.86 (-9.74-23.46)	4.23 (-10.15-18.61)
Neighbourhood (Ref = Q5)					
Q1	1.20 (1.17-1.23)	1.20 (1.18-1.23)		23.83 (4.73-42.94)	19.00 (2.75-35.25)
Q2	1.11 (1.08-1.13)	1.11 (1.08-1.14)		13.69 (-4.63-32.00)	9.45 (-6.15-25.05)
Q3	1.09 (1.06-1.12)	1.09 (1.06-1.12)		8.98 (-9.34-27.29)	7.81 (-7.79-23.41)
Q4	1.04 (1.02-1.07)	1.04 (1.02-1.07)		3.20 (-14.33-20.72)	3.60 (-11.28-18.48)
Individual + Neighbourhood					
Individual (Ref = Q5)					
Q1	1.64 (1.60-1.69)	1.32 (1.23-1.36)		47.82 (28.43-67.20)	29.50 (12.76-46.24)

Q2	1.49 (1.45-1.53)	1.20 (1.17-1.24)		35.77 (17.60-53.94)	18.00 (2.34-33.66)
Q3	1.24 (1.21-1.27)	1.11 (1.08-1.14)		16.83 (-0.49-34.14)	8.18 (-6.70-23.06)
Q4	1.09 (1.06-1.12)	1.05 (1.02-1.08)		6.17 (-10.36-22.69)	3.47 (-10.92-17.86)
Neighbourhood (Ref = Q5)					
Q1	1.11 (1.08-1.14)	1.12 (1.10-1.15)		10.29 (-8.74-29.32)	11.50 (-4.89-27.89)
Q2	1.05 (1.02-1.08)	1.06 (1.03-1.08)		4.38 (-13.72-22.48)	4.45 (-11.21-20.11)
Q3	1.03 (1.00-1.06)	1.05 (1.03-1.08)		2.42 (-15.61-20.44)	4.27 (-11.39-19.93)
Q4	1.00 (0.97-1.02)	1.02 (1.00-1.05)		-0.61 (-17.85-16.63)	1.53 (-13.35-16.41)
Individual*Neighbourhood (Ref = IQ5 and NQ5)					
IQ1 and NQ1	1.87 (1.79-1.95)	1.53 (1.47-1.60)		60.48 (56.20-64.77)	44.03 (34.65-53.40)
IQ2 and NQ2	1.62 (1.55-1.70)	1.32 (1.25-1.38)		42.56 (38.17-46.94)	24.65 (14.75-34.56)
IQ3 and NQ3	1.30 (1.23-1.37)	1.30 (1.23-1.38)		19.71 (15.42-23.99)	12.41 (2.61-22.21)
IQ4 and NQ4	1.12 (1.06-1.18)	1.09 (1.04-1.15)		7.79 (3.84-11.75)	5.64 (-3.77-15.05)

* Adjusted models control for: age, sex, tumour location, census year, province/territory at census.
* Abbreviations: HR=Hazard ratio; RD=risk difference; py=person years; Ref=reference; Q=quintile; IQ=individual quintile; NQ=neighbourhood quintile.

The overall p-value for the interaction between individual and neighbourhood income was not statistically significant (cox p=0.55), however, the individual estimates suggest some multiplicative and additive effects of individual and neighbourhood income with survival. Compared to those in individual and neighbourhood Q5, the most at risk for poor survival were those in the lowest individual and neighbourhood income quintiles (HR for IQ1+NQ1 1.87, RD 60.48). The presence of multiplicative and additive interaction for patients categorized with the same individual and neighbourhood income quintiles suggests that these two indicators are measuring different things resulting in a joint effect on survival.

After stratifying by rural residence, individual and neighbourhood income had a smaller association with survival in rural areas compared to urban areas and additive effects for neighbourhood income were not statistically significant (Supplementary Table 3). Similar patterns were observed after stratifying by stage at diagnosis (Supplementary Table 4).

DISCUSSION

Key findings

We found very weak agreement between individual and neighbourhood-level income in Canadian CRC patients, with even weaker agreement for those residing in rural areas. While both individual and neighbourhood income were associated with survival, individual income had a stronger effect on survival, with the estimate for neighbourhood income crossing the null when looking at additive effects and in rural areas. Furthermore, the presence of joint effects of individual and neighbourhood income on survival suggests that these two measures are acting independently and jointly on outcomes.

Compare to literature

The results from our study are in line with other studies demonstrating low agreement between individual and neighbourhood level income measures.[10–14,32] Other studies examining misclassification of the PCCF+ neighbourhood income found slightly stronger agreement.[10,13,14] For example, 29% and 27% perfect agreement was found by Buajitti et al and Pichora et al respectively, compared to 17% found in our study.[10,13] These studies compared individual income from the Canadian Community Health Survey (CCHS), which is entirely self-reported, to neighbourhood income from the PCCF+ which uses a combination of self-reported and tax-reported data from the long form census.[10,13] Our study used the same measures from the same populations by using the census data for both individual and neighbourhood income quintiles which could explain the difference in agreement that we observed. Moreover, our study could indicate that agreement between individual and neighbourhood income is worse in the CRC population compared to the general population.

We also demonstrated lower agreement and a slightly weaker effect of individual and neighbourhood income on survival in rural areas compared to non-rural areas. This is likely due to the larger geographical areas assigned to rural regions which result in more heterogenous populations compared to smaller urban neighbourhoods. This strong heterogeneity likely results in an attenuation of the effect on survival. Our study also presents both relative and additive effects of individual and neighbourhood income on survival where other studies present only relative effects. Reporting additive effects is especially important when examining inequalities by income because they provide information about the magnitude of inequalities in a population and are more relevant to policy makers.[33,34]

Implications

Presently, cancer organizations worldwide are acknowledging the pervasive inequalities in cancer outcomes by income and are calling for decreased barriers and equitable access to cancer care.[35–37] However, to create successful, evidence-based interventions to reduce health inequalities by income, it is imperative to have a strong understanding of the relationship between income and cancer outcomes. Using neighbourhood income as a proxy for individual income could result in an underestimation or incorrect understanding of the pathways through which income effects outcomes. This can in turn result in ineffective interventions, or even no interventions if no effect is concluded with neighbourhood income alone. Interpreting neighbourhood income as a measure of the individuals' environment is also challenging. The areas used in these measures are created from administrative units, ranging from half-block radiuses in large downtown centres to 100s of kilometers in rural areas, and do not reflect human mobility and living patterns.[38,39] In cases where researchers are interested in the effect of the

physical environment on cancer outcomes, more specific exposures can be used such as distance to cancer centres or greenspace exposure.[40,41]

Because neighbourhood income has a small effect on survival outcomes, often nearing the null, researchers should be cautious about concluding an absence of inequalities by income when individual income is not also measured. We recommend that whenever possible, individual income measures be used in cancer studies. Neighbourhood income should not be used as a proxy for individual income and should instead be interpreted as an area-level measure with the caveat that it represents different sized areas depending on rural residence. A call to data custodians to make individual measures available is required, especially for widely used cancer databases such as The Surveillance, Epidemiology, and End Results (SEER) and the National Cancer Database (NCDB).[42,43] Data confidentiality is often used as an argument for why individual measures are not provided; however, some research centres are already providing individual measures of income. For example, Statistics Canada now provides access to T1 tax files for individuals responding to the Canadian Community Health Survey.[44] For researchers that do not have access to individual measures, quantitative bias analysis using validation studies in similar populations is a relatively simple method that can be used to demonstrate the changes in effect that would occur if individual income had been available.[45]

Limitations

There are several limitations of this study. First, income was categorized into quintiles instead of measured as a continuous variable. Modelling continuous income would avoid arbitrary cut-points and increase accuracy and efficiency of the analysis. However, since most research using administrative data categorizes income as quintiles, this study demonstrates the real-world use of income. Second, individual and neighbourhood income were measured on

average 5 years prior to cancer diagnosis. This should have no effect on the analysis for agreement since both indicators are measured at the same time but might introduce misclassification bias for the effect of income on survival. For example, individual income might increase or decrease over those 5 years and neighbourhood income might change if participants moved to a different neighbourhood. Third, we did not have a large enough sample size to detect statistically significant effects for the interaction between individual and neighbourhood income, however, since specific effect sizes were statistically significant our conclusions do not change. Finally, these results may not be generalizable outside of a Canadian setting where neighbourhoods might be defined in different ways, however, the conclusion of accurately defining and interpreting income exposures remains the same.

Conclusion

Understanding differences between individual and neighbourhood income is becoming increasingly important as we shift from understanding inequalities to implementing interventions to address inequity. This study found very poor agreement between individual and neighbourhood income quintiles and much stronger associations of individual income with survival compared to neighbourhood income. Cancer researchers should avoid using neighbourhood income as a proxy for individual income, especially in CRC where patients might be at a higher risk of experiencing inequalities by income. In the absence of individual income measurements, researchers can use quantitative bias analysis to demonstrate the change in effect that might have occurred if individual income was available. Future work should use validation and quantitative bias analyses to demonstrate how neighbourhood income can be adjusted to reflect individual income.

FUNDING

This study is support by grants from the Canadian Institutes of Health Research (CIHR), the Canadian Cancer Society (MEGAN-CAN), the Canadian Centre for Applied Research in Cancer Control (ARCC) and Fonds de recherche du Québec – Santé (FRQS).

NOTES

Role of the funder. The funder did not play a role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Disclosures. The authors have no conflicts of interest.

Contribution statement. Conceptualization: LED, ALM, ECS. Data curation: LED. Formal Analysis: LED. Methodology: LED, ALM, ECS. Supervision: ALM, ECS. Writing - original draft: LED. Writing - review & editing: LED, ALM, ECS.

Prior presentations. No prior presentations.

Acknowledgements. This research was supported by funds to the Canadian Research Data Centre Network (CRDCN) from the Social Sciences and Humanities Research Council (SSHRC), the Canadian Institute for Health Research (CIHR), the Canadian Foundation for Innovation (CFI), and Statistics Canada. Although the research and analysis are based on data from Statistics Canada, the opinions expressed do not represent the views of Statistics Canada.

DATA AVAILABILITY

The CanCHEC database is protected by Statistics Canada confidentiality policies and cannot be made publicly available. Access may be granted to those who meet prespecified criteria for confidential access available at https://www.statcan.gc.ca/en/microdata/data-centres/access.

REFERENCES

1. Morgan E, Arnold M, Gini A, Lorenzoni V, Cabasag CJ, Laversanne M, et al. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. Gut [Internet]. 2022 Sep 7 [cited 2022 Nov 7]; Available from: https://gut.bmj.com/content/early/2022/09/07/gutjnl-2022-327736

2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA: A Cancer Journal for Clinicians. 2022;72(1):7–33.

3. Manser CN, Bauerfeind P. Impact of socioeconomic status on incidence, mortality, and survival of colorectal cancer patients: a systematic review. Gastrointestinal Endoscopy. 2014 Jul 1;80(1):42-60.e9.

4. Pruitt SL, Shim MJ, Mullen PD, Vernon SW, Amick BC. Association of Area Socioeconomic Status and Breast, Cervical, and Colorectal Cancer Screening: A Systematic Review. Cancer Epidemiol Biomarkers Prev. 2009 Oct 1;18(10):2579–99.

5. Sinding C, Warren R, Fitzpatrick-Lewis D, Sussman J. Research in cancer care disparities in countries with universal healthcare: mapping the field and its conceptual contours. Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer. 2014 Nov;22:3101–20.

6. Jones JRA, Berney S, Connolly B, Waterland JL, Denehy L, Griffith DM, et al. Socioeconomic Position and Health Outcomes Following Critical Illness: A Systematic Review. Critical Care Medicine. 2019 Jun;47(6):e512.

7. Afshar N, English DR, Milne RL. Factors Explaining Socio-Economic Inequalities in Cancer Survival: A Systematic Review. Cancer Control. 2021 Apr 30;28:10732748211011956.

8. Mustard CA, Derksen S, Berthelot JM, Wolfson M. Assessing ecologic proxies for household income: a comparison of household and neighbourhood level income measures in the study of population health status. Health & Place. 1999 Jun 1;5(2):157–71.

9. DeRouen MC, Schupp CW, Yang J, Koo J, Hertz A, Shariff-Marco S, et al. Impact of individual and neighborhood factors on socioeconomic disparities in localized and advanced prostate cancer risk. Cancer Causes and Control. 2018;29(10):951–66.

10. Pichora E, Polsky JY, Catley C, Perumal N, Jin J, Allin S. Comparing individual and area-based income measures: impact on analysis of inequality in smoking, obesity, and diabetes rates in Canadians 2003–2013. Can J Public Health. 2018 May 7;109(3):410–8.

11. Shariff-Marco S, Yang J, John EM, Sangaramoorthy M, Hertz A, Koo J, et al. Impact of Neighborhood and Individual Socioeconomic Status on Survival after Breast Cancer Varies by Race/Ethnicity: The Neighborhood and Breast Cancer Study. Cancer Epidemiol Biomarkers Prev. 2014 May 1;23(5):793–811.

12. DeRouen MC, Schupp CW, Koo J, Yang J, Hertz A, Shariff-Marco S, et al. Impact of individual and neighborhood factors on disparities in prostate cancer survival. Cancer Epidemiol. 2018;53:1–11.

13. Buajitti E, Chiodo S, Rosella LC. Agreement between area- and individual-level income measures in a population-based cohort: Implications for population health research. SSM Popul Health. 2020 Apr;10:100553.

14. Pampalon R, Hamel D, Gamache P. A comparison of individual and area-based socio-economic data for monitoring social inequalities in health. Health Rep. 2009 Dec;20(4):85–94.

15. Hanley GE, Morgan S. On the validity of area-based income measures to proxy household income. BMC Health Serv Res. 2008 Apr 10;8(1):79.

16. Xie S, Hubbard RA, Himes BE. Neighborhood-level measures of socioeconomic status are more correlated with individual-level measures in urban areas compared with less urban areas. Annals of Epidemiology. 2020 Mar 1;43:37-43.e4.

17. Narla NP, Pardo-Crespo MR, Beebe TJ, Sloan J, Yawn B, Williams AR, et al. Concordance between Individual vs. Area-Level Socioeconomic Measures in an Urban Setting. Journal of Health Care for the Poor and Underserved. 2015;26(4):1157–72.

18. 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. 2010 Mar 1;16(2):416–20.

19. Tjepkema M, Christidis T, Bushnik T, Pinault L. Cohort profile: The Canadian Census Health and Environment Cohorts (CanCHECs). Health Rep. 2019 Dec 18;30(12):18–26.

20. Government of Canada SC. Canadian Cancer Registry (CCR) [Internet]. 2022 [cited 2022 Nov 7]. Available from: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3207

21. Statistics Canada. Adjusted after-tax income of economic family [Internet]. 2016 [cited 2022 Apr 19]. Available from: https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=DEC&Id=103386

22. Statistics Canada. Table 18-10-0005-01 Consumer Price Index, annual average, not seasonally adjusted [Internet]. Statistics Canada; Available from: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1810000501

23. Mendelson R. Geographic Structures As Census Variables: Using Geography to Analyse Social and Economic Processes [Internet]. Ottawa: Statistics Canada; 2001 Mar [cited 2022 Jun 2]. Available from: https://www150.statcan.gc.ca/n1/pub/92f0138m/92f0138m2001001-eng.pdf

24. CIHI. Trends in Income-Related Health Inequalities in Canada [Internet]. Ottawa, Ontario: Canadian Institutes of Health Information; 2015. Available from: https://secure.cihi.ca/free_products/trends_in_income_related_inequalities_in_canada_2015_en.pdf

25. R Wilkins. PCCF + Version 3G Users Guide: Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files. Cat. No. 82F0086-XDB. 2001.

26. Dissemination area: Detailed definition [Internet]. [cited 2022 Jun 2]. Available from: https://www150.statcan.gc.ca/n1/pub/92-195-x/2011001/geo/da-ad/def-eng.htm

27. Measuring Health Inequalities: A Toolkit: Area-Level Equity Stratifiers Using PCCF and PCCF+ [Internet]. Ottawa: Canadian Institutes of Health Information; 2018. Available from: https://www.cihi.ca/sites/default/files/document/cphi-toolkit-area-level-measurement-pccf-2018-en-web.pdf

28. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol. 2010 Jun;17(6):1471–4.

29. Canadian Cancer Statistics Advisory Committee. Canadian Cancer Statistics 2018 [Internet]. Toronto, ON: Canadian Cancer Society; 2018 [cited 2018 Oct 2]. Available from: http://www.cancer.ca/~/media/cancer.ca/CW/cancer%20information/cancer%20101/Canadian%20cancer%20statistics/Canadian-Cancer-Statistics-2018-EN.pdf?la=en

30. Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychological Bulletin. 1968;70(4):213–20.

31. Martinussen T, Scheike TH. Dynamic Regression Models for Survival Data. Springer Science & Business Media; 2007. 471 p.

32. Wu CC, Hsu TW, Chang CM, Yu CH, Wang YF, Lee CC. The effect of individual and neighborhood socioeconomic status on gastric cancer survival. PloS one. 2014;9(2):e89655.

33. King NB, Harper S, Young ME. Use of relative and absolute effect measures in reporting health inequalities: structured review. BMJ [Internet]. 2012 Sep 3 [cited 2019 Dec 12];345. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3432634/

34. Mahar AL, Davis LE, Kurdyak P, Hanna TP, Coburn NG, Groome PA. Using additive and relative hazards to quantify colorectal survival inequalities for patients with a severe psychiatric illness. Annals of Epidemiology. 2021 Apr 1;56:70–4.

35. Canadian Partnership Against Cancer. Canadian Strategy for Cancer Control 2019-2029. Toronto: Canadian Partnership Against Cancer; 2018.

36. The NHS. The NHS long term plan [Internet]. London, UK; 2019 Jan. Available from: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf

37. European Commission. Europe's Beating Cancer Plan 2021-2030 [Internet]. European Commission; 2021. Available from: https://health.ec.europa.eu/system/files/2022-02/eu_cancer-plan_en_0.pdf

38. Lupton R. Neighbourhood Effects: Can We Measure Them and Does it Matter? [Internet]. Rochester, NY; 2003 [cited 2022 Oct 19]. Available from: https://papers.ssrn.com/abstract=1158964

39. Kwan MP. The Limits of the Neighborhood Effect: Contextual Uncertainties in Geographic, Environmental Health, and Social Science Research. Annals of the American Association of Geographers. 2018 Nov 2;108(6):1482–90.

40. Twohig-Bennett C, Jones A. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. Environmental Research. 2018 Oct 1;166:628–37.

41. Gulwadi GB, Joseph A, Keller AB. Exploring the Impact of the Physical Environment on Patient Outcomes in Ambulatory Care Settings. HERD. 2009 Jan 1;2(2):21–41.

42. Salem ME, Puccini A, Trufan SJ, Sha W, Kadakia KC, Hartley ML, et al. Impact of Sociodemographic Disparities and Insurance Status on Survival of Patients with Early-Onset Colorectal Cancer. Oncologist. 2021 Oct;26(10):e1730–41.

43. Dhahri A, Kaplan J, Naqvi SMH, Brownstein NC, Ntiri SO, Imanirad I, et al. The impact of socioeconomic status on survival in stage III colon cancer patients: A retrospective cohort study using the SEER census-tract dataset. Cancer Med. 2021 Aug;10(16):5643–52.

44. Canadian Cancer Registry linked to Canadian Vital Statistics Death Database and death information from the T1 Personal Master File [Internet]. Canadian Research Data Centre Network. [cited 2022 Oct 20]. Available from: https://crdcn.ca/data/canadian-cancer-registry-linked-to-canadian-vital-statistics-death-database-and-death-information-from-the-t1-personal-master-file/

45. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. Int J Epidemiol. 2014 Dec;43(6):1969–85.

3.3 Supplemental material

Supplementary Table 1. Individual and tumour characteristics by census year (N, column percent)

Variable	Whole cohort (N=103,530)	1991 (N=16,580)	1996 (N=24,835)	2001 (N=27,450)	2006 (N=23,660)	2011 (N=11,005)
Age at census (mean, SD)	65 (12)	63 (12)	64 (12)	65 (12)	66 (12)	67 (12)
Age at census						
35-44	5555 (5.4)	1125 (6.8)	1440 (5.8)	1525 (5.6)	1115 (4.7)	350 (3.2)
45-54	15240 (14.7)	2410 (14.6)	3905 (15.7)	4230 (15.4)	3385 (14.3)	1310 (11.9)
55-64	25930 (25.1)	4500 (27.2)	6240 (25.1)	6595 (24.0)	5895 (24.9)	2705 (24.6)
65-74	32195 (31.1)	5435 (32.8)	8105 (32.6)	8575 (31.2)	6770 (28.6)	3310 (30.1)
75-84	20110 (19.4)	2580 (15.6)	4300 (17.3)	5460 (19.9)	5195 (22.0)	2575 (23.4)
85+	4500 (4.6)	530 (3.2)	850 (3.4)	1075 (3.9)	1295 (5.5)	750 (6.8)
Sex						
Male	58525 (56.5)	9655 (58.2)	14035 (56.5)	15520 (56.5)	13115 (55.4)	6195 (56.3)
Female	45010 (43.5)	6925 (41.8)	10800 (43.5)	11935 (43.5)	10545 (44.6)	4805 (43.7)
Province/Territory at census						
Newfoundland and Labrador	2680 (2.6)	340 (2.1)	525 (2.1)	720 (2.6)	725 (3.1)	365 (3.3)
PEI	610 (0.6)	110 (0.7)	145 (0.6)	140 (0.5)	140 (0.6)	70 (0.7)
Nova scotia	4355 (4.2)	660 (4.0)	1025 (4.1)	1070 (2.9)	1045 (4.4)	555 (5.0)
New Brunswick	2985 (2.9)	495 (3.0)	640 (2.6)	755 (2.8)	715 (3.0)	375 (3.4)
Quebec	20345 (19.7)	4335 (26.2)	6225 (25.1)	6320 (23.0)	3445 (14.6)	15 (0.2)
Ontario	41220 (19.8)	6375 (38.5)	9585 (38.6)	10450 (38.1)	9690 (41.0)	5120 (46.5)
Manitoba	4635 (4.5)	640 (3.9)	1025 (4.1)	1235 (4.5)	1140 (4.8)	595 (5.4)
Saskatchewan	3860 (3.7)	595 (3.6)	810 (3.3)	930 (3.4)	1010 (4.3)	520 (4.7)
Alberta	8560 (8.3)	945 (5.7)	1735 (7.0)	2310 (8.4)	2315 (9.8)	1255 (11.4)
British Columbia	13900 (13.4)	2030 (12.2)	3000 (12.1)	3380 (12.3)	3390 (14.3)	2100 (19.1)
Territories combined	380 (0.4)	50 (0.3)	120 (0.5)	140 (0.5)	40 (0.2)	35 (0.3)
Rural (pop<1000)						
Urban	79000 (76.3)	12670 (76.4)	18990 (76.5)	20975 (76.4)	18000 (76.1)	8365 (76.0)
Rural	24530 (23.7)	3910 (23.6)	5850 (23.6)	6475 (23.6)	5660 (23.9)	2640 (24.0)
Tumour location						
Rectal	34810 (33.6)	5560 (33.5)	8340 (33.6)	9315 (33.9)	7900 (33.4)	3695 (33.6)
Colon	68720 (66.4)	11020 (66.5)	16500 (66.4)	18135 (66.1)	15760 (66.6)	7305 (66.4)
Status						
Alive at end of follow up	40630 (39.2)	3375 (20.3)	7215 (29.1)	11045 (40.2)	12585 (53.2)	6405 (58.2)

Died	62905 (60.8)	13205 (19.7)	17620 (70.9)	16405 (59.8)	11075 (46.8)	4600 (41.8)
Individual income quintiles						
Q1 (lowest)	23140 (22.4)	4285 (25.9)	5705 (23.0)	6775 (24.7)	4570 (19.3)	1805 (16.4)
2	23505 (22.7)	3425 (20.7)	5820 (23.4)	6270 (22.8)	5370 (22.7)	2620 (23.8)
3	19790 (19.1)	2930 (17.7)	4640 (18.7)	5185 (18.9)	4715 (19.9)	2325 (21.1)
4	18375 (17.8)	2885 (17.4)	4210 (17.0)	4670 (17.0)	4505 (19.0)	2105 (19.1)
Q5 (highest)	18715 (18.1)	3050 (18.4)	4465 (18.0)	4550 (16.6)	4495 (19.0)	2150 (19.5)
Neighbourhood income quintiles						
Q1 (lowest)	21755 (21.0)	3555 (21.5)	5265 (21.2)	5345 (19.5)	4605 (19.5)	2010 (18.3)
2	21715 (21.0)	3415 (20.6)	5080 (20.5)	5880 (21.4)	4890 (20.7)	2140 (19.5)
3	20635 (19.9)	3110 (18.8)	4905 (19.7)	5665 (20.6)	4830 (20.4)	2230 (20.3)
4	19905 (19.2)	3200 (19.3)	4775 (19.2)	5400 (19.7)	4685 (19.8)	2340 (21.3)
Q5 (highest)	19520 (18.9)	3300 (19.9)	4815 (19.4)	5160 (18.8)	4650 (19.7)	2285 (20.8)
Adjusted individual household income in 2011 Canadian dollars						
10th percentile	$17,326.80	$15,560.90	$16,107.80	$17,223.70	$18,671.29	$21,021.00
25th percentile	$25,333.90	$23,263.20	$23,563.60	$24,865.20	$27,397.00	$30,844.70
50th percentile	$41,220.50	$39,268.70	$38,195.50	$40,429.40	$44,055.34	$47,874.90

Supplementary Table 2. Agreement between individual and neighbourhood income quintiles by rural and urban residence (row percent)

Urban						
Individual income quintile	**Neighbourhood income quintile**					
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Q1	5895 (33.7)	4200 (24.0)	3100 (17.7)	2430 (13.9)	1875 (10.7)	17505
Q2	4090 (23.3)	4225 (24.1)	3705 (21.1)	3115 (17.7)	2425 (13.8)	17560
Q3	2705 (17.8)	3380 (22.3)	3195 (21.1)	3185 (21.0)	2705 (17.8)	15170
Q4	1890 (13.2)	2745 (19.2)	3030 (21.2)	3285 (23.0)	3325 (23.3)	14270
Q5	1350 (9.3)	1920 (13.3)	2580 (17.8)	3395 (23.4)	5250 (36.2)	14495
Total	15935	16470	15610	15410	15580	79000
Rural						
Individual income quintile	**Neighbourhood income quintile**					
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Q1	1585 (28.1)	1180 (20.9)	1115 (19.8)	995 (17.7)	760 (13.5)	5640
Q2	1285 (21.6)	1295 (21.8)	1260 (21.2)	1140 (19.2)	970 (16.3)	5945
Q3	815 (17.6)	955 (20.7)	1005 (21.7)	985 (21.3)	860 (18.6)	4620
Q4	655 (16.0)	795 (19.3)	875 (21.3)	935 (22.7)	845 (20.6)	4105
Q5	510 (12.0)	715 (16.9)	875 (20.7)	940 (22.2)	1185 (28.1)	4220
Total	4845	4940	5130	4995	4625	24530

Supplementary Table 3. Rural-stratified Cox proportional hazards regression and Aalen semi-parametric hazards regression for the association between income and survival

	Cox regression model				Aalen regression model			
	Rural		Urban		Rural		Urban	
Model	Unadjusted HR (95% CI)	Adjusted HR (95% CI)	Unadjusted HR (95% CI)	Adjusted HR (95% CI)	Unadjusted RD (95% CI) per 1,000 py	Adjusted RD (95% CI) per 1,000 py	Unadjusted RD (95% CI) per 1,000 py	Adjusted RD (95% CI) per 1,000 py
Individual (Ref = Q5)								
Q1	1.61 (1.53-1.70)	1.33 (1.26-1.40)	1.71 (1.67-1.76)	1.37 (1.33-1.41)	46.72 (25.40-68.04)	30.04 (11.08-49.00)	51.83 (30.94-72.72)	32.89 (14.86-50.91)
Q2	1.54 (1.46-1.62)	1.23 (1.17-1.30)	1.51 (1.46-1.55)	1.22 (1.19-1.26)	40.88 (20.56-61.20)	22.16 (4.41-39.90)	36.35 (16.54-56.17)	18.91 (1.74-36.08)
Q3	1.23 (1.67-1.31)	1.10 (1.04-1.16)	1.26 (1.22-1.30)	1.13 (1.10-1.17)	17.26 (-2.41-36.94)	7.96 (-9.21-25.13)	18.21 (-0.46-36.89)	9.93 (-6.17-26.02)
Q4	1.10 (1.04-1.17)	1.07 (1.01-1.14)	1.10 (1.06-1.13)	1.06 (1.02-1.09)	7.45 (-11.94-26.83)	6.28 (-10.82-23.38)	6.72 (-11.24-24.67)	3.80 (-11.87-19.46)
Neighbourhood (Ref = Q5)								
Q1	1.15 (1.09-1.21)	1.10 (1.05-1.16)	1.31 (1.28-1.35)	1.24 (1.20-1.27)	14.16 (-7.59-35.91)	10.18 (-8.70-29.07)	26.68 (5.79-47.57)	21.46 (3.72-39.20)
Q2	1.08 (1.03-1.14)	1.03 (0.98-1.09)	1.18 (1.15-1.22)	1.13 (1.01-1.16)	7.81 (-12.94-28.56)	3.56 (-14.40-21.52)	15.33 (-4.42-35.07)	11.13 (-5.68-27.94)
Q3	1.10 (1.04-1.15)	1.06 (1.00-1.11)	1.11 (1.07-1.14)	1.10 (1.07-1.13)	9.05 (-12.05-30.16)	6.24 (-11.93-24.41)	8.72 (-10.74-28.18)	8.14 (-8.38-24.67)
Q4	1.03 (0.98-1.09)	1.01 (0.96-1.06)	1.04 (1.01-1.07)	1.05 (1.02-1.08)	3.14 (-17.68-23.96)	1.52 (-16.58-19.62)	3.06 (-15.76-21.87)	4.01 (-12.01-20.04)
Individual + Neighbourhood								
Individual (Ref = Q5)								
Q1	1.60 (1.52-1.69)	1.32 (1.25-1.39)	1.66 (1.61-1.71)	1.32 (1.28-1.36)	45.99 (24.46-67.52)	29.27 (10.10-48.45)	48.55 (27.30-69.79)	29.35 (10.96-47.73)
Q2	1.53 (1.45-1.61)	1.22 (1.16-1.29)	1.47 (1.43-1.52)	1.19 (1.16-1.23)	40.51 (20.05-60.98)	21.68 (3.80-39.57)	34.09 (14.06-54.12)	16.46 (-0.92-33.85)
Q3	1.23 (1.16-1.30)	1.09 (1.03-1.16)	1.24 (1.20-1.28)	1.11 (1.08-1.15)	16.90 (-2.85-36.64)	7.70 (-9.54-24.94)	16.71 (-2.17-35.60)	8.21 (-8.03-24.45)
Q4	1.10 (1.04-1.17)	1.07 (1.01-1.14)	1.09 (1.05-1.12)	1.05 (1.01-1.08)	7.15 (-12.23-26.54)	6.06 (-11.11-23.23)	5.88 (-12.15-23.90)	2.80 (-12.94-18.54)
Neighbourhood (Ref = Q5)								
Q1	1.05 (1.00-1.11)	1.05 (1.00-1.11)	1.12 (1.09-1.16)	1.15 (1.11-1.18)	5.07 (-16.75-26.89)	4.82 (-14.35-23.99)	11.97 (-6.70-30.64)	13.43 (-4.74-31.60)
Q2	1.03 (0.97-1.08)	1.01 (0.96-1.06)	1.06 (1.03-1.09)	1.07 (1.04-1.10)	2.48 (-18.19-23.16)	0.63 (-17.40-18.66)	5.00 (-14.39-24.39)	5.69 (-11.40-22.79)
Q3	1.05 (1.00-1.11)	1.04 (0.99-1.09)	1.02 (0.99-1.05)	1.06 (1.03-1.09)	5.07 (-15.96-26.11)	4.12 (-14.26-22.51)	1.60 (-18.22-21.42)	4.34 (-12.40-21.08)
Q4	1.01 (0.96-1.06)	1.00 (0.95-1.05)	0.99 (0.96-1.02)	1.03 (1.00-1.06)	0.65 (-20.03-21.32)	0.14 (-18.03-18.31)	-0.98 (-22.16-20.19)	1.88 (-14.22-17.98)
Individual*Neighbourhood (Ref = IQ5 and NQ5)								
IQ1 and NQ1	1.67 (1.51-1.84)	1.38 (1.25-1.53)	1.92 (1.83-2.02)	1.57 (1.49-1.65)	48.88 (39.50-58.25)	32.81 (23.43-42.18)	63.65 (54.27-73.02)	46.81 (37.43-56.18)
IQ2 and NQ2	1.61 (1.45-1.79)	1.27 (1.14-1.41)	1.62 (1.53-1.71)	1.32 (1.25-1.39)	44.88 (34.98-54.79)	24.63 (14.73-34.54)	41.64 (31.80-51.54)	24.19 (14.28-34.09)
IQ3 and NQ3	1.34 (1.19-1.50)	1.16 (1.04-1.30)	1.28 (1.20-1.36)	1.18 (1.11-1.25)	24.12 (14.32-33.93)	12.84 (3.03-22.64)	18.15 (8.34-27.95)	11.92 (2.12-21.73)

IQ4 and NQ4	1.09 (0.96-1.23)	1.10 (0.98-1.24)	1.11 (1.04-1.18)	1.08 (1.02-1.15)	10.71 (1.29-20.12)	8.54 (-0.87-17.96)	6.88 (-2.54-16.29)	4.74 (-4.68-14.15)

* Adjusted models control for: age, sex, tumour location, census year, province/territory at census

* Abbreviations: HR=Hazard ratio; CI=confidence interval; RD=risk difference; py=person years; Ref=reference; Q=quintile; IQ=individual quintile; NQ=neighbourhood quintile.

Supplementary Table 4. Stage stratified Cox proportional hazards regression for the association between income and survival. Cohort limited to patients diagnosed in 2010 or after and who had complete stage information.

	Stage I	**Stage II**	**Stage III**	**Stage IV**
Model	**HR (95% CI)**	**HR (95% CI)**	**HR (95% CI)**	**HR (95% CI)**
Individual no covariates (Ref = Q5)				
Q1	2.74 (2.23-3.38)	2.04 (1.74-2.39)	2.18 (1.91-2.48)	1.42 (1.29-1.57)
Q2	2.17 (1.77-2.67)	1.69 (1.45-1.98)	1.84 (1.62-2.09)	1.42 (1.29-1.56)
Q3	1.93 (1.56-2.39)	1.36 (1.15-1.60)	1.47 (1.29-1.67)	1.24 (1.13-1.37)
Q4	1.46 (1.17-1.82)	1.08 (0.90-1.28)	1.25 (1.09-1.44)	1.15 (1.04-1.27)
Individual with covariates (Ref = Q5)				
Q1	1.84 (1.48-2.28)	1.53 (1.30-1.80)	1.73 (1.52-1.97)	1.26 (1.14-1.39)
Q2	1.29 (1.04-1.59)	1.17 (1.00-1.38)	1.40 (1.23-1.59)	1.24 (1.13-1.37)
Q3	1.35 (1.09-1.67)	1.08 (0.91-1.27)	1.24 (1.09-1.41)	1.15 (1.04-1.27)
Q4	1.21 (0.97-1.51)	0.99 (0.83-1.18)	1.20 (1.05-1.37)	1.11 (0.997-1.22)
Neighbourhood no covariates (Ref = Q5)				
Q1	1.58 (1.31-1.91)	1.38 (1.19-1.61)	1.31 (1.16-1.48)	1.20 (1.09-1.32)
Q2	1.26 (1.04-1.52)	1.30 (1.17-1.50)	1.23 (1.09-1.39)	1.09 (0.99-1.21)
Q3	1.21 (1.00-1.47)	1.12 (0.96-1.30)	1.14 (1.01-1.29)	1.04 (0.95-1.15)
Q4	1.05 (0.87-1.27)	1.00 (0.85-1.17)	1.08 (0.95-1.22)	1.05 (0.96-1.16)
Neighbourhood with covariates (Ref = Q5)				
Q1	1.47 (1.21-1.77)	1.39 (1.20-1.61)	1.24 (1.09-1.40)	1.19 (1.07-1.31)
Q2	1.12 (0.92-1.36)	1.22 (1.05-1.42)	1.22 (1.08-1.38)	1.08 (0.98-1.19)
Q3	1.16 (0.96-1.40)	1.13 (0.97-1.32)	1.14 (1.01-1.29)	1.05 (0.96-1.16)
Q4	1.04 (0.85-1.26)	0.97 (0.83-1.14)	1.05 (0.93-1.19)	1.08 (1.00-1.19)
Individual + Neighbourhood no covariates (Ref = Q5)				
Individual				
Q1	2.58 (2.08-3.20)	1.94 (1.65-2.29)	2.13 (1.87-2.43)	1.39 (1.25-1.54)
Q2	2.09 (1.67-2.58)	1.63 (1.39-1.92)	1.81 (1.59-2.06)	1.40 (1.27-1.54)
Q3	1.88 (1.52-2.32)	1.32 (1.12-1.56)	1.45 (1.27-1.65)	1.23 (1.11-1.36)
Q4	1.44 (1.15-1.80)	1.07 (0.89-1.27)	1.24 (1.08-1.42)	1.14 (1.03-1.26)
Neighbourhood				
Q1	1.25 (1.03-1.52)	1.13 (0.97-1.32)	1.07 (0.94-1.21)	1.10 (0.99-1.21)
Q2	1.07 (0.88-1.30)	1.14 (0.98-1.32)	1.09 (0.96-1.23)	1.03 (0.94-1.14)
Q3	1.07 (0.88-1.30)	1.02 (0.87-1.19)	1.03 (0.91-1.17)	1.01 (0.91-1.11)
Q4	0.99 (0.82-1.20)	0.95 (0.81-1.11)	1.02 (0.90-1.15)	1.03 (0.93-1.13)
Individual + Neighbourhood with covariates (Ref = Q5)				
Individual				
Q1	1.72 (1.38-2.14)	1.41 (1.89-1.67)	1.68 (1.47-1.92)	1.23 (1.10-1.36)
Q2	1.24 (1.01-1.54)	1.10 (0.94-1.30)	1.37 (1.20-1.56)	1.22 (1.11-1.35)
Q3	1.32 (1.07-1.64)	1.04 (0.87-1.23)	1.22 (1.06-1.39)	1.14 (1.03-1.26)
Q4	1.20 (0.96-1.51)	0.97 (0.81-1.15)	1.19 (1.04-1.36)	1.10 (0.99-1.21)
Neighbourhood				
Q1	1.32 (1.08-1.60)	1.23 (1.08-1.47)	1.09 (0.96-1.24)	1.13 (1.02-1.25)
Q2	1.06 (0.87-1.28)	1.16 (1.00-1.35)	1.13 (1.00-1.28)	1.05 (0.95-1.15)
Q3	1.09 (0.90-1.33)	1.09 (0.94-1.28)	1.08 (0.95-1.22)	1.03 (0.94-1.14)
Q4	1.01 (0.83-1.23)	0.95 (0.81-1.12)	1.01 (0.89-1.14)	1.06 (0.96-1.17)
Individual*Neighbourhood no covariates (Ref = IQ5*NQ5)				

IQ1 and NQ1	3.32 (2.37-4.67)	2.20 (1.69-2.86)	2.63 (2.09-3.30)	1.51 (1.27-1.79)
IQ2 and NQ2	2.00 (1.39-2.89)	1.91 (1.45-2.52)	2.14 (1.68-2.72)	1.49 (1.25-1.79)
IQ3 and NQ3	1.94 (1.33-2.84)	1.09 (0.78-1.52)	1.58 (1.22-2.04)	1.21 (1.00-1.47)
IQ4 and NQ4	1.32 (0.88-1.98)	0.95 (0.68-1.31)	1.33 (1.02-1.74)	1.28 (1.05-1.56)
Individual*Neighbourhood with covariates (Ref = IQ5*NQ5)				
IQ1 and NQ1	2.20 (1.56-3.11)	1.86 (1.42-2.42)	2.15 (1.71-2.70)	1.38 (1.16-1.64)
IQ2 and NQ2	1.08 (0.74-1.56)	1.41 (1.07-1.85)	1.77 (1.39-2.25)	1.31 (1.09-1.56)
IQ3 and NQ3	1.29 (0.88-1.89)	0.95 (0.68-1.32)	1.44 (1.12-1.87)	1.12 (0.92-1.36)
IQ4 and NQ4	1.06 (0.71-1.59)	0.88 (0.64-1.23)	1.31 (1.00-1.70)	1.27 (1.04-1.54)

* Adjusted models control for: age, sex, tumour location, census year, province/territory at census

* Abbreviations: HR=Hazard ratio; CI=confidence interval; Ref=reference; Q=quintile; IQ=individual quintile; NQ=neighbourhood quintile.

Chapter 4 : Manuscript 2. Probabilistic bias analysis for household income exposure misclassification by neighbourhood in a cohort of colorectal cancer patients

4.1 Preface

Although neighbourhood income is a poor proxy for individual income, it is still consistently used in health studies due to the lack of access to individual income measures in routinely collected data in Canada, the US, the UK, and Australia. I wanted to provide a solution to estimate the effects of individual income in the absence of neighbourhood income until a time when individual income or quantiles based on individual income are available in all routinely collected databases. I was in a unique situation where I had both individual and neighbourhood income for all individuals; therefore, I treated neighbourhood income as a misclassification problem for individual income and demonstrated how to use probabilistic bias analysis to provide bias-adjusted estimates of the effect of neighborhood income on 5-year mortality to approximate the effect of individual income more closely. I hope that future research can extend this work to different populations and outcomes to provide bias-adjusted effects of income when only neighbourhood income is available.

This manuscript is in the process of being submitted to the *International Journal of Epidemiology.*

Citation: Davis LE, Banack HR, Calderon-Anyosa R, Strumpf EC, Mahar AL. Probabilistic bias analysis for household income exposure misclassification by neighbourhood in a cohort of colorectal cancer patients. Submit to the International Journal of Epidemiology.

4.2 Manuscript 2

ABSTRACT

Introduction. Despite poor agreement, neighbourhood income is used as a proxy for household income due to a lack of data availability. We quantified misclassification between household and neighbourhood income and demonstrate quantitative bias analysis (QBA) in scenarios where only neighbourhood income is available in assessing income inequalities on colorectal cancer mortality.

Methods. This was a retrospective study of adults with colorectal cancer diagnosed 2006-2014 from Statistics Canada's Canadian Census Health and Environment Cohort. Neighbourhood income quintiles from Statistics Canada were used. Census, household income quintiles were used to determine bias parameters and confirm results of the QBA. We calculated positive and negative predictive values using multinomial models, adjusting for rural residence, age and sex. Probabilistic QBA was conducted to explore the implication of exposure misclassification when estimating the effect of income on 5-year mortality.

Results. We found poor agreement between neighbourhood and household income: positive predictive values ranged from 21% to 37%. The bias-adjusted risk of neighbourhood income on 5-year mortality was similar to the risk of mortality by household income. The bias-adjusted relative risk of the lowest income quintile compared to the highest was 1.42 (95% SI 1.32-1.53) compared to 1.46 (95% CI 1.39-1.54) for household income and 1.18 (95% CI 1.12-1.24) for neighbourhood income.

Conclusion. QBA can be used to estimate adjusted effects of neighbourhood income on mortality that represent household income. The predictive values from our study can be applied

to similar cohorts with only neighbourhood income to estimate the effects of household income on cancer mortality.

KEYWORDS. Area-level income, neighbourhood income, individual income, household income, quantitative bias analysis, exposure misclassification, income inequalities, cancer, colorectal cancer

INTRODUCTION

Income, whether measured at the individual, household, or area level, is a key social determinant of health, affecting outcomes ranging from health behaviours to diabetes to cancer.[1,2] Individual income, often measured as household income, will affect the individual's ability to access material goods and services. Area-level income refers to the average or median income in a census tract, dissemination area, or other geographic region, like a neighbourhood. While both measures are valid, without the measurement of both individual and neighbourhood income in a multi-level model, it is impossible to understand the effect of neighbourhood income as an individual's physical environment (e.g., walkability or green space).[3,4] Unfortunately, due to limitations on data availability, health services researchers and social epidemiologists using routinely collected data often use neighbourhood income as a proxy for household-level income.[5–7] Studies investigating cancer, health behaviours, chronic disease, and child health across Canada, the US and the UK have all demonstrated that neighbourhood income is a poor proxy for household income, with poor agreement and less variation in neighbourhood than household income.[3,4,6,8–10] Consequently, using neighbourhood income as a substitute for household income may underestimate the impact of income on health outcomes and can lead to incorrect conclusions regarding income inequalities.[3,4,11–14]

Quantitative bias analysis is a simple but underused method to obtain quantitative estimates of the direction, magnitude and uncertainty arising from exposure misclassification.[15–17] To date, this method has rarely been used in scenarios with polytomous exposures to quantify misclassification at each level of a categorical variable.[18] This paper describes an approach to address the common practice of using neighbourhood income quintile as a proxy for household income quintile in a five by five scenario when evaluating income as an exposure. In a cohort of

colorectal cancer (CRC) patients we aimed to 1) To investigate the potential magnitude of exposure misclassification when neighbourhood income quintile is used instead of household income quintile and 2) To demonstrate the application of quantitative bias analysis in scenarios where only neighbourhood income quintile is available in assessing inequalities in the risk of mortality.

METHODS

Study design and population

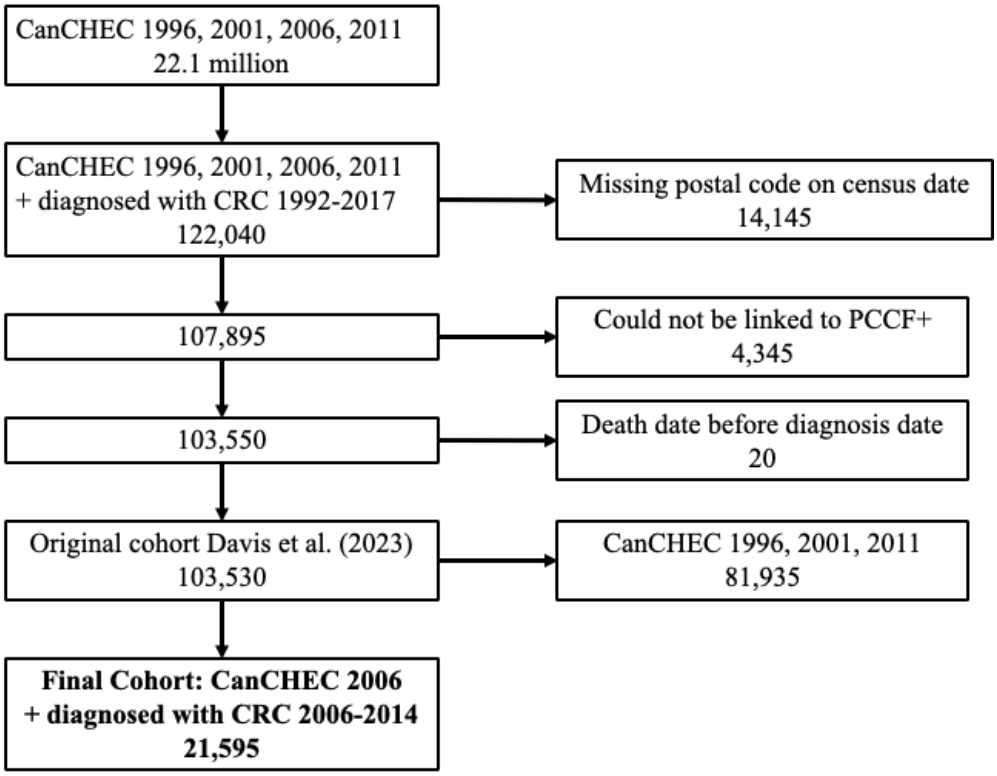
This was a retrospective cohort study using the Canadian Census Health and Environment Cohorts (CanCHEC).[3,19] CanCHEC is a nationwide cohort created by Statistics Canada, which integrates census data, vital statistics, and cancer registry databases. This cohort enables the follow-up of individuals who participate in the long-form census, allowing for the analysis of mortality and cancer-related outcomes.[19] The census is performed every 5 years in Canada, and approximately 1 in every 5 households completes the long-form census. Individuals are eligible to complete the long-form census if they were residents of Canada on census day and were not living in an institution, such as a nursing home, penitentiary, or group home. The Canadian Cancer Registry collects cancer data on all Canadians diagnosed with cancer, except Quebec residents after 2010.[20] The postal code conversation file (PCCF+) was linked to the CanCHEC using the participants' postal code at census to obtain neighbourhood income quintiles and rural residence.

We used a subset of individuals with CRC from the CanCHEC created for a previous study.[3] Briefly, the previous study included individuals over the age of 35 with a new diagnosis

of CRC (International Classification of Diseases for Oncology (ICD-O-3) codes: C180, C182-C189, C199, C209) between January 1, 1992 and December 31, 2017. Included individuals completed at least one long-form census in the ten years prior to their cancer diagnosis (census cycles 1991-2011). Individuals with a missing date of diagnosis, missing postal code at census, postal code that could not be linked to the PCCF+ and those with a death date before their diagnosis date were excluded. For this study, we further limited the cohort to individuals who responded to the 2006 census and had a CRC diagnosis between January 1, 2006, and December 31, 2014 (Figure 1). The 2006 census was the first census to use tax files to measure household income, and this period allowed for 5 years of complete follow-up on all individuals.

Ethics approval was obtained from the McGill University Research Ethics Board (#A04-M37-22A) and Statistics Canada.

FIGURE 4.1 COHORT CREATION



*Abbreviations: CanCHEC = Canadian Census Health and Environment Cohorts; CRC = Colorectal cancer; PCCF+ = Postal code conversation file plus

We had household and neighbourhood income quintile measures for all individuals in the cohort. Household income was defined as the true exposure and neighbourhood income as misclassified household income. Both household and neighbourhood income were measured at the time and residence of the 2006 census and represented income before the cancer diagnosis.

Household income quintile was created to be as similar as possible to neighborhood income and defined as adjusted before-tax household income using the long-form census. All sources of income from the calendar year before the census (2005) were summed for each household and adjusted for the number of household members using the single-person

equivalence scales from Statistics Canada's low-income cut-offs.[21] In 2006, individuals responding to the census had the option to consent to income tax linkage. All individuals included in the cohort consented to the use of income tax linkage. Household income quintiles were created by ordering individual household incomes within each census metropolitan area (CMA), census agglomeration (CA) and other regions by province within the full weighted census cohort, as opposed to the CRC cohort. We then divided each category into 5 equal groups to create quintiles specific for each CMA/CA/other region. CMAs are large urban areas of ≥100,000 people, CAs are smaller areas of ≥10,000 population, and other regions incorporate urban fringe and rural areas.[22] CMA/CA specific quintiles take into account differences in cost of living across regions.

Neighbourhood income quintiles were created by Statistics Canada for the PCCF+ and are the most widely used area-level income measure in Canada.[23] The PCCF+ uses Statistics Canada Census Profile Data at the dissemination area (DA) level to calculate area-level adjusted household income.[24] DAs are Statistics Canada's smallest geographical area representing approximately 400-700 individuals per area.[25] Total income for each DA is calculated by multiplying the median, before-tax income of that area by the total number of households and adjusted for household size by dividing by the sum of the single-person equivalents of the DA to obtain the median household income per single-person equivalent for each DA. CMA/CA quintiles are then constructed by ranking DAs within each CMA/CA/other region by province from lowest to highest, then dividing into fifths.[26] We linked the PCCF+ to the CanCHEC using the postal code in 2006.

Outcome

Mortality was defined as death from any cause from the vital statistics database occurring within 5 years from diagnosis. We followed everyone in our cohort for a maximum of 5 years from the date of the cancer diagnosis or until death.

Covariates

Predictors of misclassification of household income by neighbourhood income quintiles were conceptualized as rural residence, age and sex. Previous work has demonstrated that larger rural areas suffer from greater misclassification compared to smaller urban areas.[3] Rural residence was defined according to the PCCF+ as residing in a census subdivision with a population of <1,000 and a population density of <400 persons per square kilometre. Continuous age and dichotomous sex were defined according to the census. For the bias analysis, we controlled for the same variables: rural residence, continuous age and sex. We stratified results by province/territory at the time of the census to obtain bias parameters that could be applicable across Canada or by province/territory. Due to small sample sizes, we grouped Prairie provinces (Alberta, Manitoba, and Saskatchewan), and Atlantic provinces (New Brunswick, Nova Scotia, Prince Edward Island, and Newfoundland) and excluded territories for stratified analyses. Additionally, to describe the cohort, we examined tumour location (colon or rectum) and stage at diagnosis, which were obtained from the CCR.[27] Stage was largely missing prior to 2010, as this was when provinces and territories started prioritizing routine collection of stage data for lung, colorectal, breast and prostate cancers.[28]

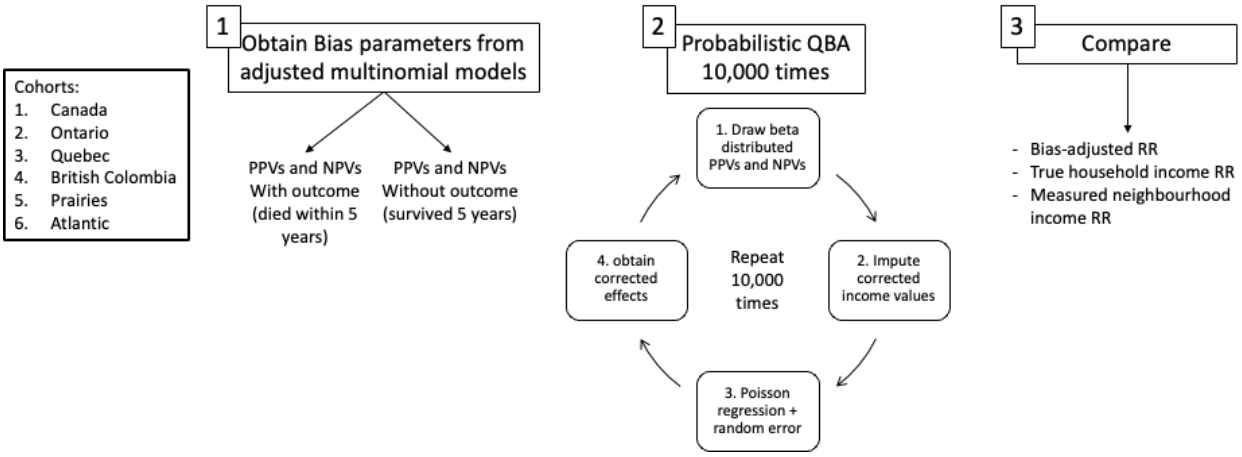
Statistical Analysis

The analysis involved three main steps. First, we calculated the bias parameters (PPV and NPV) to determine the probability of being truly classified in household income quintiles 1-5

given the observed classification of neighbourhood income quintiles 1-5. We calculated the bias parameters separately for those who survived and did not survive 5 years.[29] Second, we performed probabilistic quantitative bias analyses to obtain bias-adjusted measures of the effect of neighbourhood income on 5-year mortality. Third, we compared the bias-adjusted estimates to the estimates of the effect of household income on 5-year mortality (considered correct) and the misclassified neighbourhood income estimates (considered incorrect). We performed each of these steps within the Canadian CRC cohort and stratified by province of residence. These steps are defined in more detail in Figure 2 and below.

R Studio version 4.2.2 was used for all analyses. In keeping with Statistics Canada's data confidentiality guidelines, counts were rounded to the nearest 5.

FIGURE 4.2. ANALYSIS STEPS



*Abbreviations: PPVs = Positive predictive values; NPVs = Negative predictive values; QBA = quantitative bias analysis; RR = relative risk

Step one: Obtain bias parameters

Bias parameters were obtained for each strata of the outcome. Traditionally, negative and positive predictive values are used in dichotomous scenarios, however, we extended these

methods to calculate bias parameters at each combination of quintiles (i.e. 25 values total).[17] Positive predictive values (PPVs) represent the probability of being correctly classified in household income quintiles 1-5 given the same classification by neighbourhood income quintiles (5 values). Negative predictive values (NPVs) indicate the probability of being truly classified in household income quintiles 1-5, despite an incorrect classification of observed neighbourhood income quintile for each of the remaining four quintiles (totalling 20 values). For example, the NPVs for income quintile 1 would be calculated based on the classification in neighbourhood income quintiles 2, 3, 4, and 5 for a total of 4 NPVs. Supplemental Table 1 provides an example of how to calculate crude predictive values from a 5 by 5 table.

Extending previous methods for dichotomous scenarios, we calculated negative and positive predictive values from multinomial logistic regression models, for those who survived 5 years and those who did not.[30,31] The use of multinomial models allows for the adjustment of predictors of misclassification which were included as rural residence, age and sex.[30,31] To calculate PPVs and NPVs using this model, we defined the true exposure (household income quintile) as the outcome and the misclassified exposure (neighbourhood income quintile) as the predictor for those who died within 5 years and those who survived.[17,30] Using the model coefficients, we predicted the probability that household income quintile corresponded to neighbourhood income quintile while holding all covariates at their mean. To do this, we used the R package "Effects" to obtain the predictive values and their 95% confidence intervals at each of the 25 levels of household and neighbourhood income quintiles to obtain a 5 by 5 matrix of predictive values.[32] Adjusted PPVs and NPVs were calculated for the whole Canadian cohort and by province (Quebec, Ontario, British Colombia, Prairie provinces and Atlantic provinces).

Step two: Probabilistic quantitative bias analysis

We used the PPVs and NPVs from step one to conduct a probabilistic bias analysis with Monte Carlo sampling to adjust for exposure misclassification.[17,18,31] We performed the bias analysis in the whole Canadian cohort and stratified by province. The bias analysis consisted of record-level correction for exposure misclassification and estimating the effect of income on 5-year mortality using the bias-adjusted exposure variable.

In Step 2.1 we used a beta distribution to incorporate uncertainty in the predictive values obtained in Step 1.[17,31] The beta distribution can model a wide range of probability density shapes, which is ideal for proportions, and does not yield values outside of an allowed range, such as with the normal distribution.[17] We parameterize the beta distribution by choosing upper and lower values of the predictive values that are likely with only 2.5% probability of seeing a value lower than the low end and 2.5% probability of seeing a value higher than the high end of the range.[17] The ends of the range (U and L) and the predictive values (x) are used in the following equations to obtain alpha and beta values.[17] We defined U and L as the lower and upper values of the 95% confidence intervals from the adjusted PPVs and NPVs.

$$sd = \frac{U - L}{2 \cdot 1.96}$$

$$\alpha = x\left(\frac{x(1-x)}{sd^2} - 1\right)$$

$$\beta = (1-x)\left(\frac{x(1-x)}{sd^2} - 1\right)$$

In step 2.2, we imputed the corrected income quintiles for each combination of neighbourhood income quintiles and the outcome for each observation in the dataset (10 possible

combinations). We used the predictive values obtained from a random draw from the beta distribution calculated in step 2.1 to reassign each neighbourhood income quintile into the bias-adjusted income quintile. For example, an individual observed to be in neighbourhood income quintile 1 who died within 5 years would be reassigned based on the PPV of being classified as household income quintile 1 and based on the NPV of being classified in household income quintile 2-5 and having the outcome (died). This step creates a new dataset with the imputed bias-adjusted values for income.

In step 2.3, we used the bias-adjusted dataset to estimate the association between corrected neighbourhood income and 5-year mortality. We used Poisson regression with robust error variance to obtain relative risks.[33,34] Robust variance accounts for the potential overestimation of errors that can occur when applying Poisson regression to binomial data.[34] We present unadjusted and adjusted models for age, sex and rural residence. In step 2.4 we accounted for total study error by incorporating random error into each of the bias-adjusted effect estimates by adding a number drawn randomly from a standard normal distribution.[17]

Steps 2.1 to 2.3 were repeated 10,000 times to create a distribution of relative risks. The bias-adjusted effect estimates reported are the 50th percentile of the distribution. The 2.5th and 97.5th percentile of the distribution provides a 95% simulation interval (SI) around the bias-adjusted estimate.

Step three: Effect of individual and neighbourhood income on 5-year mortality

Since we had true household income for all individuals in the cohort, we were able to compare the bias-adjusted estimated effects of neighbourhood income to the true estimated effects using household income. We used unadjusted and adjusted Poisson regression with robust

error variance to obtain relative risks and 95% confidence intervals (CI) for the effects of household and neighbourhood income on 5-year mortality. Adjusted models included age, sex, and rural residence.

RESULTS

The cohort included 21,595 CRC patients diagnosed between 2006 and 2014. Cohort characteristics are described in Table 1. The mean age was 66, 55.3% were male, 23.8% lived in rural areas and most patients resided in Ontario (40.5%) followed by Quebec (15.8%). 42.4% of patients died within 5 years.

TABLE 4.1 COHORT CHARACTERISTICS (N=21,595)

Variables	N (%)
Age at diagnosis (mean (SD))	65.95 (12.46)
Sex	
Male	11945 (55.3)
Female	9650 (44.7)
Tumour location	
Rectal	7245 (33.6)
Colon	14350 (66.4)
Rural residence	
Not rural	16460 (76.2)
Rural	5135 (23.8)
Province at diagnosis	
Newfoundland and Labrador	650 (3.0)
PEI	125 (0.6)
Nova scotia	945 (4.4)
New Brunswick	630 (2.9)
Quebec	3415 (15.8)
Ontario	8750 (40.5)
Manitoba	1020 (4.7)
Saskatchewan	910 (4.2)

Alberta	2055 (9.5)
British Columbia	3055 (14.1)
Territories	40 (0.2)
Diagnosis year	
2006	2620 (12.1)
2007	2750 (12.7)
2008	2960 (13.7)
2009	2890 (13.4)
2010	3000 (13.9)
2011	1800 (8.3)
2012	1870 (8.7)
2013	1895 (8.8)
2014	1810 (8.4)
Stage at diagnosis	
0	80 (0.4)
I	2035 (9.4)
II	2400 (11.1)
III	2580 (11.9)
IV	1790 (8.3)
Unknown	460 (2.1)
Missing	12255 (56.7)
Died within 5 years	
No	12450 (56.6)
Yes	9150 (42.4)
Neighbourhood income quintiles	
1	4215 (19.5)
2	4490 (20.8)
3	4400 (20.4)
4	4250 (19.7)
5	4245 (19.6)
Household income quintiles	
1	4240 (19.6)
2	4940 (22.9)
3	4305 (19.9)
4	4070 (18.8)
5	4050 (18.7)

The PPVs and NPVs for Canada are presented in Table 2 and by province in Supplemental Table 3. We presented adjusted predictive values to retain precision, but crude and adjusted values were similar. Overall, PPVs were very low, ranging from 20.64% (95% CI 18.75-22.68) to 36.93% (95% CI 35.05-38.84). Compared with those who died within 5 years, individuals who survived 5 years had higher PPVs at the highest income quintile (36.93% vs 27.69%) and lower values at the lowest income quintile (31.27% vs 36.62%). The highest PPV by province was for those residing in the Prairie provinces, correctly classified in the lowest income quintile among those who died within 5 years (42.8%) (Supplemental Table 3).

TABLE 4.2. ADJUSTED PREDICTIVE VALUES AND 95% CONFIDENCE INTERVALS STRATIFIED BY 5-YEAR MORTALITY FOR ALL OF CANADA

	Household income quintiles									
		Count					Adjusted predictive values (95% confidence interval)				
Neighbourhood income quintile	**Overall**									
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**
	Q1	1440	1130	735	555	360	33.64 (32.20-35.11)	27.04 (25.70-28.42)	17.84 (16.70-19.05)	13.22 (12.22-14.29)	8.26 (7.46-9.13)
	Q2	955	1195	945	805	590	20.68 (19.51-21.90)	26.48 (25.19-27.81)	21.65 (20.45-22.89)	18.22 (17.10-19.39)	12.93 (12.02-14.00)
	Q3	795	1065	935	870	735	17.63 (16.52-18.80)	24.27 (23.01-25.57)	21.84 (20.63-23.10)	20.00 (18.83-21.22)	16.26 (15.19-17.39)
	Q4	600	850	920	940	940	13.94 (12.92-15.02)	20.18 (18.99-21.44)	22.29 (21.05-23.59)	22.28 (21.04-23.57)	21.30 (20.09-22.57)
	Q5	450	705	770	895	1425	10.23 (9.35-11.18)	16.43 (15.34-17.59)	18.69 (17.53-19.91)	21.51 (20.29-22.79)	33.14 (31.72-34.59)
	Survived 5 years (N=12,447)										
		Q1	**Q2**	**Q3**	**Q4**	**Q5**					
	Q1	710	590	405	325	205	31.27 (29.36-33.26)	26.51 (24.70-28.41)	18.57 (17.00-20.26)	14.65 (13.23-16.19)	8.99 (7.87-10.25)
	Q2	455	615	560	490	395	17.62 (16.17-19.17)	24.14 (22.49-25.87)	22.78 (21.17-24.49)	19.84 (18.30-21.46)	15.62 (14.24-17.11)
	Q3	415	575	560	540	470	15.78 (14.41-17.25)	22.28 (20.70-23.96)	22.49 (20.89-24.16)	21.32 (19.76-22.97)	18.13 (16.68-19.69)
	Q4	310	465	575	600	635	11.90 (10.70-13.22)	17.99 (16.54-19.54)	22.76 (21.16-24.43)	23.36 (21.75-25.05)	24.00 (22.37-25.70)
	Q5	220	370	455	560	950	8.36 (7.35-9.50)	14.27 (12.96-15.69)	18.19 (16.73-19.75)	22.25 (20.66-23.92)	36.93 (35.05-38.84)
	Died within 5 years (N=9,150)										
		Q1	**Q2**	**Q3**	**Q4**	**Q5**					
	Q1	730	540	330	230	155	36.62 (34.48-38.80)	27.56 (25.61-29.60)	16.83 (15.23-18.57)	11.52 (10.17-13.01)	7.47 (6.40-8.72)
	Q2	500	575	385	315	195	24.80 (22.93-26.78)	29.45 (27.46-31.52)	20.01 (18.28-21.85)	16.04 (14.47-17.74)	9.70 (8.47-11.09)
	Q3	380	490	370	330	260	20.30 (18.51-22.23)	26.91 (24.90-29.01)	20.76 (18.94-22.70)	18.16 (16.44-20.01)	13.88 (12.36-15.54)
	Q4	290	385	345	345	305	17.08 (15.33-18.98)	23.34 (21.34-25.46)	21.37 (19.44-23.43)	20.64 (18.75-22.68)	17.57 (15.84-19.48)
	Q5	230	335	315	335	470	13.12 (11.58-14.82)	19.65 (17.80-21.63)	19.25 (17.41-21.22)	20.30 (18.43-22.31)	27.69 (25.57-29.91)

*predictive values and confidence intervals obtained from multinomial models

*adjusted predictive values for rural residence, age and sex

Household income had a greater estimated effect on 5-year mortality compared to neighbourhood income, in all of Canada and across provinces (Table 3). For example, patients experiencing the lowest household income had a 46% (95% CI 1.39-1.54) increase in the risk of 5-year mortality compared to those experiencing the highest income, while those living in neighbourhoods with the lowest income had only an 18% (95% CI 1.12-1.24) increased risk of 5-year mortality compared to those living in neighbourhoods with the highest income. Adjusting for age, sex and rural residence resulted in attenuation of the effect of household income (RR 1.26 (95%CI 1.20-1.33)) but not neighbourhood income (RR 1.14 (95% CI 1.09-1.19).

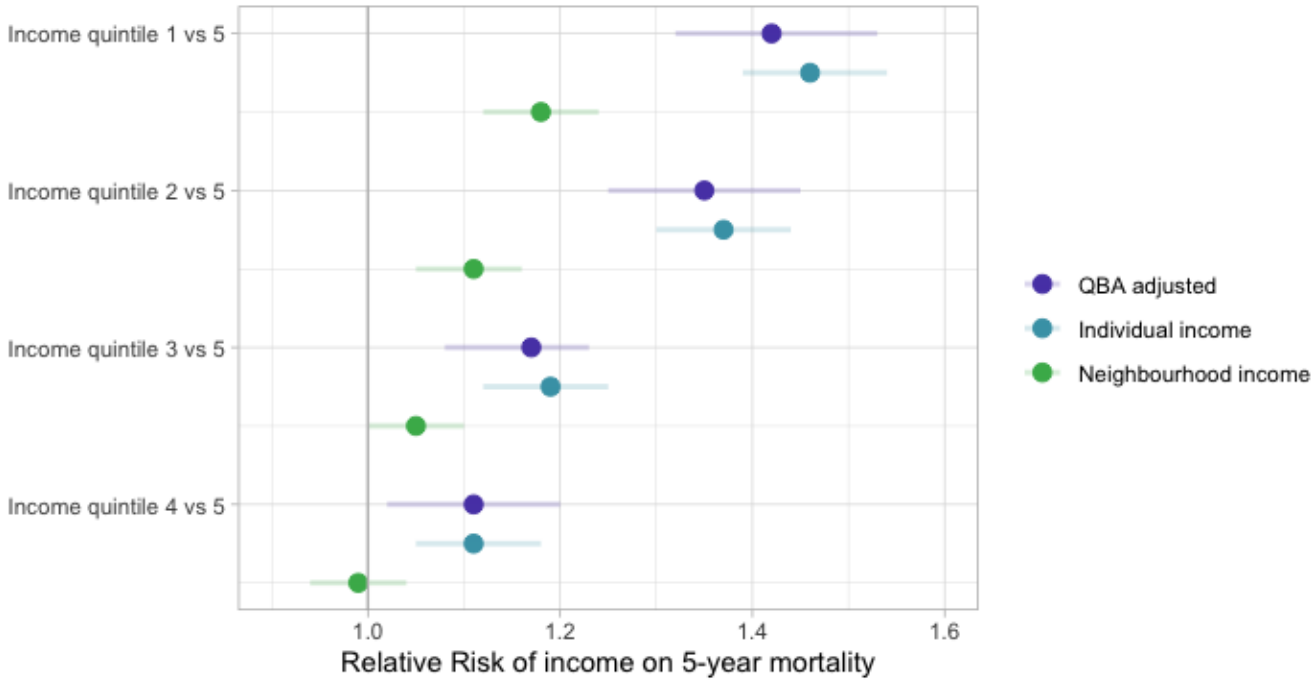
The results of the bias analysis are presented in Table 3 and Figure 1. After accounting for both random and systematic error, the bias-adjusted 5-year mortality risk was similar to the risk of mortality by household income, with overlapping confidence intervals. For example, the bias-adjusted relative risk of being in the lowest income quintile compared to the highest for Canada was 1.42 (95% SI 1.32-1.53) compared to 1.46 (95% CI 1.39-1.54) for household income and 1.18 (95% CI 1.12-1.24) for neighbourhood income. After adjusting for age, sex, and rural residence, bias-adjusted effects of income were greater than the adjusted effect of household income on 5-year mortality. This pattern persisted across provinces, with all provinces demonstrating similar bias-adjusted effects compared to household income.

TABLE 4.3. BIAS ADJUSTED RELATIVE RISKS OF DEATH WITHIN FIVE YEARS (YES/NO) WITH SYSTEMATIC AND RANDOM ERROR AND 95% SIMULATION INTERVALS FOR ALL OF CANADA AND BY PROVINCE, REFERENCE CATEGORY IS THE HIGHEST INCOME QUINTILE FOR ALL COMPARISONS (QUINTILE 5)

Scenario	Simulation RR with random error (95% SI)	True household income RR (95% CI)	Measured neighbourhood income RR (95%CI)
Canada			
Quintile 1	1.42 (1.32-1.53)	1.46 (1.39-1.54)	1.18 (1.12-1.24)
Quintile 2	1.35 (1.25-1.45)	1.37 (1.30-1.44)	1.11 (1.05-1.16)
Quintile 3	1.17 (1.08-1.28)	1.19 (1.12-1.25)	1.05 (1.00-1.10)
Quintile 4	1.11 (1.02-1.20)	1.11 (1.05-1.18)	0.99 (0.94-1.04)
Ontario			
Quintile 1	1.39 (1.23-1.55)	1.42 (1.31-1.54)	1.15 (1.06-1.24)
Quintile 2	1.32 (1.18-1.48)	1.34 (1.24-1.45)	1.10 (1.02-1.19)
Quintile 3	1.11 (0.95-1.24)	1.10 (1.01-1.21)	1.03 (0.95-1.12)
Quintile 4	1.07 (0.95-1.19)	1.06 (0.97-1.16)	0.95 (0.88-1.04)
Quebec			
Quintile 1	1.35 (1.13-1.69)	1.42 (1.25-1.61)	1.19 (1.05-1.34)
Quintile 2	1.31 (1.11-1.62)	1.35 (1.19-1.53)	1.13 (1.00-1.28)
Quintile 3	1.24 (1.01-1.53)	1.28 (1.12-1.47)	1.03 (0.91-1.17)
Quintile 4	1.11 (0.88-1.41)	1.11 (0.96-1.28)	1.06 (0.93-1.21)
British Colombia			
Quintile 1	1.40 (1.13-1.71)	1.41 (1.23-1.63)	1.18 (1.04-1.34)
Quintile 2	1.34 (1.09-1.63)	1.36 (1.18-1.56)	1.09 (0.96-1.25)
Quintile 3	1.29 (1.04-1.60)	1.31 (1.13-1.51)	0.99 (0.86-1.14)
Quintile 4	1.19 (0.96-1.47)	1.21 (1.04-1.40)	1.02 (0.89-1.17)
Prairie provinces			
Quintile 1	1.52 (1.27-1.83)	1.59 (1.40-1.80)	1.28 (1.14-1.44)
Quintile 2	1.48 (1.24-1.78)	1.48 (1.30-1.68)	1.16 (1.03-1.30)
Quintile 3	1.20 (0.97-1.46)	1.18 (1.03-1.35)	1.15 (1.02-1.30)
Quintile 4	1.06 (0.84-1.31)	1.06 (0.91-1.22)	1.03 (0.91-1.18)
Atlantic provinces			
Quintile 1	1.49 (1.22-1.87)	1.54 (1.32-1.81)	1.10 (0.95-1.27)
Quintile 2	1.36 (1.13-1.74)	1.39 (1.18-1.63)	1.03 (0.89-1.19)
Quintile 3	1.25 (0.89-1.49)	1.24 (1.05-1.47)	1.05 (0.91-1.21)
Quintile 4	1.25 (0.99-1.61)	1.30 (1.10-1.54)	0.89 (0.76-1.04)

*****Abbreviations: RR = relative risk; CI = confidence interval; SI = simulation interval

FIGURE 4.3. COMPARING BIAS ADJUSTED RELATIVE RISK TO THE RELATIVE RISK OF HOUSEHOLD AND NEIGHBOURHOOD INCOME ON 5-YEAR MORTALITY FOR ALL OF CANADA



DISCUSSION

Using routinely collected data from a representative sample of Canadians with CRC, we provide detailed methodology on how to use quantitative bias analysis to estimate bias-adjusted effects of neighbourhood income on 5-year mortality. By comparing bias-adjusted estimates to the estimated effect of household income and measured neighbourhood income on 5-year mortality, we were able to demonstrate that the quantitative bias analysis provided similar effects to household income. Moreover, this is one of few studies detailing how to obtain bias parameters from multinomial models for multi-categorical variables and apply them to quantitative bias analysis.[18]

Our study is in line with previous studies that have also found poor agreement between individual or household and neighbourhood income quintiles, with the proportion categorized as

the same individual and neighbourhood income ranging from 34% to 37%.[3,4,35] Other solutions to the problem of limited access to individual-level income variables have been proposed. For example, researchers have developed a housing-based socioeconomic index (HOUSES) that correlated strongly with other socioeconomic factors, such as individual income and education levels.[36,37] This method has been used to describe socioeconomic inequalities in asthma research, vaccine research, and child and youth health outcomes.[37–39] Another study in the US developed a method of approximating individual-level income using income probabilities from the US census.[40] However, both these methods require access to census data or other individual-level housing information, which is not always available. Our method allows researchers to use the PPVs and NPVs presented in our study to obtain bias-adjusted effects of income on 5-year mortality in cohorts with similar prevalence of exposure, covariates and predictors of misclassification.

Implications

With increasing calls to understanding socioeconomic inequities in the healthcare system, and cancer care specifically, it is imperative that studies examine both household and neighbourhood-level income inequalities.[41–43] By treating the measurement of income as a misclassification problem, our study provides one solution to the pervasive lack of individual-level income variables in routinely collected or secondary data sources. The predictive values from our study can be applied directly to perform quantitative bias analyses in CRC cohorts with similar prevalence of exposure, covariates and predictors of misclassification to determine the effect of household income on 5-year mortality, when only neighbourhood income is available. Moreover, we provide detailed methodology and code that can be replicated in different disease

sites and other countries to obtain bias parameters that can be applied to studies examining a wide range of disease contexts and outcomes using income inequalities as the main exposure. For example, future research might aim to broaden the scope of this study by examining PPVs and NPVs in the whole Canadian population and applying them to different outcomes or cohorts. While researchers wait for data custodians to provide access to individual-level income measurements, including quintiles, our methods can be built upon to provide bias-adjusted estimates of the effects of household income on outcomes.

Strengths

Our study contributes methodologically by extending quantitative bias analysis to a scenario that is not normally considered exposure misclassification. We used a measurement of neighbourhood income that is broadly used across Canada and present a method for understanding the effects of household income on colorectal cancer mortality when only neighbourhood income is available. By having both household and neighbourhood income for all individuals in the cohort, we were able to demonstrate the validity of using quantitative bias analysis to provide an adjusted estimate of the effect of income on mortality. This is also one of the few studies that demonstrate the use of quantitative bias analysis for a multi-categorical exposure.

Limitations

Our study has some limitations. First, since we used predictive values, which need to be calculated separately for each outcome, therefore our bias parameters can only be applied to studies examining income and mortality. Moreover, for simplicity, we examined 5-year mortality instead of survival. Since many cancer studies examine survival, this may limit the

generalizability of our work. Second, the predictive values from our validation study can only be applied to external cohorts that have a similar prevalence of exposure, covariates and predictors of misclassification; therefore, our results are not generalizable outside of Canada or even to other cancer sites. More work needs to be done to calculate predictive values in other cancer sites or to determine if predictive values for all Canadians can be applied to subgroups such as patients with cancer. Finally, we used a subset of CRC patients who responded to the 2006 census. While the census is supposed to be a representative sample of the population, it excludes those in institutions such as prisons and care homes, where having low income may be more likely, resulting in a cohort with fewer patients experiencing low income. However, we compared our cohort to patients diagnosed within the same timeframe in the CCR, and the cohorts were similar (Supplemental Table 4).

Conclusion

In an ideal world, all administrative databases would include both area-level and individual-level socioeconomic variables, such as individual or household income quintiles from tax records. Until then, our study provides a relatively simple method to estimate the effect of household income on cancer mortality when only neighbourhood income is available. Importantly, we provide foundational methodological processes for future studies to replicate our work in other diseases sites and countries where income is the primary exposure of interest.

Author statement

LE Davis: Conceptualization, methodology, formal analysis, data curation, writing – original draft, visualization, funding acquisition. **HR Banack:** Methodology, writing – review and editing. **R Calderon-Anyosa:** Methodology, validation, writing – review and editing. **EC Strumpf:** Conceptualization, methodology, writing – review and editing, supervision, funding acquisition. **AL Mahar:** Conceptualization, methodology, writing – review and editing, supervision, funding acquisition.

Funding

This study is support by grants from the Canadian Institutes of Health Research (CIHR), the Canadian Cancer Society (MEGAN-CAN), the Canadian Centre for Applied Research in Cancer Control (ARCC) and Fonds de recherche du Québec – Santé (FRQS).

Role of the funder. The funder did not play a role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Disclosures. The authors have no conflicts of interest.

Prior presentations. No prior presentations.

Acknowledgements. This research was supported by funds to the Canadian Research Data Centre Network (CRDCN) from the Social Sciences and Humanities Research Council (SSHRC), the Canadian Institute for Health Research (CIHR), the Canadian Foundation for Innovation (CFI), and Statistics Canada. Although the research and analysis are based on data from Statistics Canada, the opinions expressed do not represent the views of Statistics Canada.

DATA AVAILABILITY

The CanCHEC database is protected by Statistics Canada confidentiality policies and cannot be made publicly available. Access may be granted to those who meet prespecified criteria for confidential access available at https://www.statcan.gc.ca/en/microdata/data-centres/access.

REFERENCES

1. Kawachi I, Kennedy BP. Socioeconomic determinants of health : Health and social cohesion: why care about income inequality? BMJ. 1997 Apr 5;314(7086):1037.

2. Marmot MG. Understanding Social Inequalities in Health. Perspect Biol Med. 2003 Jul 24;46(3):S9–23.

3. Davis LE, Mahar AL, Strumpf EC. Agreement between individual and neighborhood income measures in patients with colorectal cancer in Canada. JNCI J Natl Cancer Inst. 2023 Jan 27;djad017.

4. Buajitti E, Chiodo S, Rosella LC. Agreement between area- and individual-level income measures in a population-based cohort: Implications for population health research. SSM - Popul Health. 2020 Apr 1;10:100553.

5. Moss JL, Johnson NJ, Yu M, Altekruse SF, Cronin KA. Comparisons of individual- and area-level socioeconomic status as proxies for individual-level measures: evidence from the Mortality Disparities in American Communities study. Popul Health Metr. 2021 Jan 7;19(1):1.

6. Ingleby FC, Belot A, Atherton I, Baker M, Elliss-Brookes L, Woods LM. Assessment of the concordance between individual-level and area-level measures of socio-economic deprivation in a cancer patient cohort in England and Wales. BMJ Open. 2020 Nov 26;10(11):e041714.

7. Afshar N, English DR, Milne RL. Factors Explaining Socio-Economic Inequalities in Cancer Survival: A Systematic Review. Cancer Control J Moffitt Cancer Cent. 2021 Apr 30;28:10732748211011956.

8. Narla NP, Pardo-Crespo MR, Beebe TJ, Sloan J, Yawn B, Williams AR, et al. Concordance between Individual vs. Area-Level Socioeconomic Measures in an Urban Setting. J Health Care Poor Underserved. 2015;26(4):1157–72.

9. Xie S, Hubbard RA, Himes BE. Neighborhood-level measures of socioeconomic status are more correlated with individual-level measures in urban areas compared with less urban areas. Ann Epidemiol. 2020 Mar 1;43:37-43.e4.

10. Jerrim J. Measuring socio-economic background using administrative data. What is the best proxy available? DoQSS Work Pap [Internet]. 2020 Aug 1 [cited 2023 Jul 11]; Available from: https://ideas.repec.org//p/qss/dqsswp/2009.html

11. Booth CM, Li G, Zhang-Salomons J, Mackillop WJ. The impact of socioeconomic status on stage of cancer at diagnosis and survival: a population-based study in Ontario, Canada. Cancer. 2010 Sep 1;116:4160–7.

12. Berglund A, Garmo H, Robinson D, Tishelman C, Holmberg L, Bratt O, et al. Differences according to socioeconomic status in the management and mortality in men with high risk prostate cancer. Eur J Cancer. 2012;48(1):75–84.

13. Jembere N, Campitelli MA, Sherman M, Feld JJ, Lou W, Peacock S, et al. Influence of Socioeconomic Status on Survival of Hepatocellular Carcinoma in the Ontario Population; A Population-Based Study, 1990–2009. PLoS ONE. 2012 Jul 13;7(7):e40917.

14. Johnson AM, Hines RB, Johnson JA, Bayakly AR. Treatment and survival disparities in lung cancer: the effect of social environment and place of residence. Lung Cancer Amst Neth. 2014 Mar;83(3):401–7.

15. Petersen JM, Ranker LR, Barnard-Mayers R, MacLehose RF, Fox MP. A systematic review of quantitative bias analysis applied to epidemiological research. Int J Epidemiol. 2021 Oct 1;50(5):1708–30.

16. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. Int J Epidemiol. 2014 Dec;43(6):1969–85.

17. Fox MP, MacLehose RF, Lash TL. Applying Quantitative Bias Analysis to Epidemiologic Data. Springer Nature; 2022. 475 p.

18. Radin RG, Rothman KJ, Hatch EE, Mikkelsen EM, Sorensen HT, Riis AH, et al. Maternal recall error in retrospectively-reported time-to-pregnancy: an assessment and bias analysis. Paediatr Perinat Epidemiol. 2015 Nov;29(6):576–88.

19. Tjepkema M, Christidis T, Bushnik T, Pinault L. Cohort profile: The Canadian Census Health and Environment Cohorts (CanCHECs). Health Rep. 2019 Dec 18;30(12):18–26.

20. Government of Canada SC. Canadian Cancer Registry (CCR) [Internet]. 2022 [cited 2022 Nov 7]. Available from: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3207

21. Statistics Canada. Adjusted after-tax income of economic family [Internet]. 2016 [cited 2022 Apr 19]. Available from: https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=DEC&Id=103386

22. Mendelson R. Geographic Structures As Census Variables: Using Geography to Analyse Social and Economic Processes [Internet]. Ottawa: Statistics Canada; 2001 Mar [cited 2022 Jun 2]. Available from: https://www150.statcan.gc.ca/n1/pub/92f0138m/92f0138m2001001-eng.pdf

23. CIHI. Trends in Income-Related Health Inequalities in Canada [Internet]. Ottawa, Ontario: Canadian Institutes of Health Information; 2015. Available from: https://secure.cihi.ca/free_products/trends_in_income_related_inequalities_in_canada_2015_en.pdf

24. R Wilkins. PCCF + Version 3G Users Guide: Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files. Cat. No. 82F0086-XDB. 2001.

25. Dissemination area: Detailed definition [Internet]. [cited 2022 Jun 2]. Available from: https://www150.statcan.gc.ca/n1/pub/92-195-x/2011001/geo/da-ad/def-eng.htm

26. Measuring Health Inequalities: A Toolkit: Area-Level Equity Stratifiers Using PCCF and PCCF+ [Internet]. Ottawa: Canadian Institutes of Health Information; 2018. Available from: https://www.cihi.ca/sites/default/files/document/cphi-toolkit-area-level-measurement-pccf-2018-en-web.pdf

27. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol. 2010 Jun;17(6):1471–4.

28. Canadian Cancer Statistics Advisory Committee. Canadian Cancer Statistics 2018 [Internet]. Toronto, ON: Canadian Cancer Society; 2018 [cited 2018 Oct 2]. Available from: http://www.cancer.ca/~/media/cancer.ca/CW/cancer%20information/cancer%20101/Canadian%20cancer%20statistics/Canadian-Cancer-Statistics-2018-EN.pdf?la=en

29. Fox MP, Lash TL, Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. Int J Epidemiol. 2005 Dec 1;34(6):1370–6.

30. Coughlin SS, Trock B, Criqui MH, Pickle LW, Browner D, Tefft MC. The logistic modeling of sensitivity, specificity, and predictive value of a diagnostic test. J Clin Epidemiol. 1992 Jan;45(1):1–7.

31. Banack HR, Stokes A, Fox MP, Hovey KM, Cespedes-Feliciano EM, LeBlanc E, et al. Stratified probabilistic bias analysis for BMI-related exposure misclassification in postmenopausal women. Epidemiol Camb Mass. 2018 Sep;29(5):604–13.

32. Fox J, Andersen R. Effect Displays for Multinomial and Proportional-Odds Logit Models. Sociol Methodol. 2006 Aug 1;36(1):225–55.

33. Chen W, Qian L, Shi J, Franklin M. Comparing performance between log-binomial and robust Poisson regression models for estimating risk ratios under model misspecification. BMC Med Res Methodol. 2018 Jun 22;18(1):63.

34. Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J Epidemiol. 2004 Apr 1;159(7):702–6.

35. Pichora E, Polsky JY, Catley C, Perumal N, Jin J, Allin S. Comparing individual and area-based income measures: impact on analysis of inequality in smoking, obesity, and diabetes rates in Canadians 2003–2013. Can J Public Health Rev Can Santé Publique. 2018 May 7;109(3):410–8.

36. Juhn YJ, Beebe TJ, Finnie DM, Sloan J, Wheeler PH, Yawn B, et al. Development and Initial Testing of a New Socioeconomic Status Measure Based on Housing Data. J Urban Health. 2011 Oct 1;88(5):933–44.

37. Harris MN, Lundien MC, Finnie DM, Williams AR, Beebe TJ, Sloan JA, et al. Application of a novel socioeconomic measure using individual housing data in asthma research: an exploratory study. Npj Prim Care Respir Med. 2014 Jun 26;24(1):1–6.

38. Bjur KA, Wi CI, Ryu E, Derauf C, Crow SS, King KS, et al. Socioeconomic Status, Race/Ethnicity, and Health Disparities in Children and Adolescents in a Mixed Rural-Urban Community—Olmsted County, Minnesota. Mayo Clin Proc. 2019 Jan 1;94(1):44–53.

39. Hammer R, Capili C, Wi CI, Ryu E, Rand-Weaver J, Juhn YJ. A new socioeconomic status measure for vaccine research in children using individual housing data: a population-based case-control study. BMC Public Health. 2016 Sep 21;16(1):1000.

40. Kim U, Koroukian SM, Stange KC, Spilsbury JC, Dong W, Rose J. Describing and assessing a new method of approximating categorical individual-level income using community-level income from the census (weighting by income probabilities). Health Serv Res. 2022;57(6):1348–60.

41. Canadian Partnership Against Cancer. Canadian Strategy for Cancer Control 2019-2029. Toronto: Canadian Partnership Against Cancer; 2018.

42. The NHS. The NHS long term plan [Internet]. London, UK; 2019 Jan. Available from: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf

43. European Commission. Europe's Beating Cancer Plan 2021-2030 [Internet]. European Commission; 2021. Available from: https://health.ec.europa.eu/system/files/2022-02/eu_cancer-plan_en_0.pdf

4.3 Supplemental material

Supplemental Table 1. Details on how to calculate crude PPVs and NPVs

		Household income quintiles					
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Neighbourhood income quintile total**
Neighbourhood income quintiles	**Q1**	**PPV1** 1440/4220 = 0.34	**NPV1.2** 1130/4220 = 0.27	**NPV1.3** 735/4220 = 0.17	**NPV1.4** 555/4220 = 0.13	**NPV1.5** 360/4220 = 0.09	4220
	Q2	**NPV2.1** 955/4490 = 0.21	**PPV2** 1195/4490 = 0.27	**NPV2.3** 945/4490 = 0.21	**NPV2.4** 805/4490 = 0.18	**NPV2.5** 590/4490 = 0.13	4490
	Q3	**NPV3.1** 795/4400 = 0.18	**NPV3.2** 1065/4400 = 0.24	**PPV3** 935/4400 = 0.21	**NPV3.4** 870/4400 = 0.20	**NPV3.5** 735/4400 = 0.17	4400
	Q4	**NPV4.1** 600/4250 = 0.14	**NPV4.2** 850/4250 = 0.20	**NPV4.3** 920/4250 = 0.22	**PPV4** 940/4250 = 0.22	**NPV4.5** 940/4250 = 0.22	4250
	Q5	**NPV5.1** 450/4245 = 0.11	**NPV5.2** 705/4245 = 0.17	**NPV5.3** 770/4245 = 0.18	**NPV5.4** 895/4245 = 0.21	**PPV5** 1425/4245 = 0.34	4245

Supplemental Table 2. Cohort characteristics for colorectal cancer patients diagnosed 2001-2005 in the Canadian Census Health and Environment Cohorts (N=13,515)

	N (%)
Age at diagnosis (mean (SD))	66.82 (11.82)
Sex	
Male	7610 (56.3)
Female	5905 (43.7)
Tumour location	
Rectal	4615 (34.2)
Colon	8900 (65.8)
Rural residence	
Not rural	10280 (76.0)
Rural	3240 (24.0)
Province at diagnosis	
Newfoundland and Labrador	320 (2.4)
Prince Edward Island	75 (0.6)
Nova scotia	550 (4.1)
New Brunswick	380 (2.8)
Quebec	3315 (24.5)
Ontario	5060 (37.4)
Manitoba	590 (4.4)
Saskatchewan	440 (3.3)
Alberta	1050 (7.8)
British Columbia	1640 (12.1)
Territories	100 (0.7)
Diagnosis year	
2001	2445 (18.1)
2002	2635 (19.5)
2003	2680 (19.8)
2004	2765 (20.5)
2005	2995 (22.2)
Died within 5 years	
No	4570 (33.8)
Yes	8945 (66.2)
Neighbourhood income quintiles	
1	2645 (19.6)
2	2980 (22.0)
3	2845 (21.0)
4	2565 (19.0)
5	2485 (18.4)

Individual income quintiles	
1	3535 (26.2)
2	3170 (23.5)
3	2505 (18.5)
4	2190 (16.2)
5	2115 (15.6)

Supplemental Table 3. Province stratified adjusted predictive values and 95% confidence intervals

*predictive values and confidence intervals obtained from multinomial models

*adjusted predictive values for rural residence, age and sex (NB. The difference between crude and adjusted values were very small (<1%), therefore crude values are not presented)

	Individual income quintiles																			
		Count					**Adjusted predictive values (95% CI)**													
Canada																				
							Predictive value	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**
Neighbourhood income quintiles	**Overall (N=21,595)**																			
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	1440	1130	735	555	360	30.81	26.81	35.11	23.93	20.31	27.97	20.40	17.01	24.27	13.46	10.67	16.83	11.40	8.85	14.58
	Q2	955	1195	945	805	590	20.69	17.27	24.59	25.09	21.38	29.20	22.44	18.91	26.42	19.39	16.08	23.20	12.38	9.74	15.62
	Q3	795	1065	935	870	735	18.41	15.21	22.12	24.01	20.39	28.05	22.77	19.22	26.76	19.18	15.89	22.98	15.62	12.64	19.15
	Q4	600	850	920	940	940	17.74	14.39	21.66	19.64	16.13	23.69	17.25	13.95	21.16	22.35	18.64	26.57	23.02	19.25	27.27
	Q5	450	705	770	895	1425	13.46	10.53	17.05	17.90	14.54	21.83	19.69	16.20	23.73	19.29	15.83	23.29	29.66	25.50	34.20
	Survived 5 years (N=12,445)																			
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	710	590	405	325	205	28.54	23.26	34.47	24.51	19.56	30.24	19.43	14.96	24.85	14.42	10.57	19.37	13.11	9.44	17.92
	Q2	455	615	560	490	395	17.63	13.49	22.71	24.59	19.76	30.16	24.22	19.42	29.77	20.41	15.96	25.72	13.14	9.59	17.75
	Q3	415	575	560	540	470	15.77	11.86	20.65	23.32	18.58	28.85	25.15	20.23	30.80	18.23	13.98	23.43	17.53	13.37	22.64
	Q4	310	465	575	600	635	12.73	9.22	17.32	18.09	13.88	23.24	19.09	14.77	24.31	23.07	18.36	28.56	27.03	21.98	32.74
	Q5	220	370	455	560	950	11.84	8.31	16.59	13.70	9.91	18.63	18.45	14.08	23.82	17.70	13.42	22.98	38.31	32.38	44.61
	Died within 5 years (N=9,150)																			
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	730	540	330	230	155	33.46	27.62	39.86	23.16	18.14	29.08	21.42	16.57	27.22	12.23	8.59	17.13	9.72	6.51	14.27
	Q2	500	575	385	315	195	24.39	19.02	30.69	25.61	20.15	31.95	20.34	15.43	26.31	17.86	13.27	23.60	11.81	8.14	16.83
	Q3	380	490	370	330	260	21.71	16.74	27.66	24.74	19.45	30.92	19.91	15.12	25.76	19.99	15.19	25.83	13.66	9.72	18.86

	Q4	290	385	345	345	305	25.42	19.34	32.63	21.74	16.10	28.68	14.16	9.59	20.40	21.27	15.63	28.25	17.42	12.33	24.04
	Q5	230	335	315	335	470	13.12	11.58	14.82	19.65	17.80	21.63	19.25	17.41	21.22	20.30	18.43	22.31	27.69	25.57	29.91
Quebec																					
Neighbourhood income quintiles							**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**
	Overall (N=3,415)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	275	195	100	80	70	36.43	32.91	40.11	27.40	24.22	30.84	14.76	12.31	17.61	11.52	9.34	14.11	9.88	7.87	12.34
	Q2	175	205	150	115	85	22.86	19.94	26.06	28.24	25.06	31.65	20.96	18.12	24.11	15.94	13.42	18.81	12.01	9.82	14.61
	Q3	140	180	135	135	135	19.15	16.41	22.23	25.17	22.11	28.51	18.77	16.07	21.82	18.83	16.12	21.88	18.07	15.42	21.07
	Q4	85	135	125	125	145	14.27	11.68	17.32	22.76	19.56	26.31	20.26	17.24	23.66	20.31	17.29	23.70	22.40	19.26	25.89
	Q5	75	100	120	145	185	11.89	9.55	14.71	16.19	13.46	19.34	19.35	16.41	22.69	23.05	19.88	26.55	29.52	26.03	33.27
	Survived 5 years (N=1,890)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	135	95	50	45	40	35.04	30.23	40.17	26.49	22.16	31.32	14.70	11.38	18.79	12.54	9.48	16.42	11.23	8.35	14.96
	Q2	85	105	80	65	60	20.34	16.62	24.64	26.32	22.15	30.97	21.20	17.37	25.62	16.65	13.22	20.76	15.48	12.17	19.50
	Q3	70	95	75	90	85	16.94	13.60	20.91	22.97	19.13	27.32	18.22	14.77	22.26	21.49	17.78	25.74	20.38	16.76	24.56
	Q4	45	75	60	70	95	12.46	9.35	16.41	22.41	18.25	27.19	18.26	14.49	22.74	20.43	16.47	25.04	26.45	22.03	31.40
	Q5	35	50	70	95	120	9.11	6.55	6.55	14.32	11.06	11.06	19.48	15.70	15.70	25.39	21.16	21.16	31.71	27.10	27.10
	Died within 5 years (N=1,525)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	140	100	50	35	30	37.84	32.79	43.17	28.60	24.03	33.66	14.59	11.21	18.77	10.40	7.57	14.12	8.57	6.01	12.09
	Q2	90	100	70	50	30	25.80	21.40	30.75	30.30	25.61	35.44	20.58	16.58	25.25	15.15	11.69	19.40	8.17	5.68	11.61
	Q3	70	85	60	45	45	22.23	17.86	27.30	27.93	23.14	33.28	19.56	15.45	24.45	15.24	11.60	19.78	15.04	11.43	19.53
	Q4	45	60	60	55	50	16.57	12.50	21.63	22.82	18.12	28.32	22.73	18.07	28.18	20.35	15.92	25.64	17.53	13.42	22.56
	Q5	40	45	50	50	70	15.86	11.84	20.92	18.78	14.40	24.11	19.03	14.64	24.36	19.68	15.22	25.06	26.65	21.55	32.47
Ontario																					
							Predictive value	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**

Neighbourhood income quintiles	**Overall (N=8,750)**																			
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	525	470	280	220	125	32.38	30.11	34.74	29.33	27.13	31.63	17.61	15.81	19.56	13.41	11.83	15.17	7.27	6.11	8.62
	Q2	370	470	375	335	265	19.83	18.04	21.74	25.70	23.71	27.78	21.20	19.36	23.17	18.87	17.11	20.76	14.41	12.85	16.13
	Q3	295	420	370	370	295	16.41	14.73	18.24	24.00	22.04	26.09	21.69	19.80	23.71	21.36	19.48	23.37	16.53	14.85	18.37
	Q4	245	375	425	400	385	13.17	11.68	14.82	20.68	18.86	22.63	23.85	21.93	25.89	21.99	20.13	23.97	20.31	18.51	22.23
	Q5	165	300	305	370	605	9.14	7.88	10.58	16.96	15.26	18.81	18.04	16.29	19.94	21.55	19.66	23.57	34.30	32.07	36.60
	Survived 5 years (N=5,130)																			
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	265	235	175	140	65	29.97	27.00	33.13	27.11	24.24	30.18	19.97	17.43	22.77	15.69	13.42	18.27	7.26	5.73	9.15
	Q2	175	240	225	205	175	16.75	14.56	19.18	23.02	20.52	25.73	22.71	20.22	25.42	20.34	17.95	22.96	17.18	14.96	19.65
	Q3	150	235	230	230	185	14.27	12.25	16.55	22.71	20.23	25.39	22.61	20.14	25.29	22.61	20.14	25.29	17.80	15.57	20.28
	Q4	135	215	275	260	260	11.49	9.75	13.49	18.73	16.54	21.13	24.62	22.17	27.24	22.94	20.56	25.50	22.23	19.89	24.77
	Q5	80	165	180	230	400	7.27	5.86	8.99	15.12	13.07	17.42	17.34	15.15	19.77	22.34	19.90	24.98	37.93	35.01	40.95
	Died within 5 years (N=3,630)																			
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	260	235	110	80	55	35.33	31.90	38.92	32.03	28.72	35.54	14.65	12.27	17.41	10.66	8.62	13.10	7.33	5.67	9.42
	Q2	195	230	150	130	90	24.06	21.19	27.19	29.31	26.21	32.61	18.98	16.37	21.89	16.80	14.33	19.60	10.85	8.86	13.22
	Q3	145	185	140	135	105	19.58	16.82	22.68	25.92	22.81	29.29	20.11	17.31	23.24	19.53	16.76	22.64	14.86	12.41	17.69
	Q4	110	160	150	140	130	15.92	13.35	18.87	23.69	20.63	27.06	22.35	19.36	25.65	20.52	17.64	23.73	17.52	14.85	20.55
	Q5	85	140	130	135	205	12.11	9.88	14.75	19.84	17.02	23.01	18.87	16.10	22.00	20.24	17.38	23.44	28.94	25.63	32.48
British Colombia																					
Neighbourhood income quintiles							**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**
	Overall (N=3,055)																			
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	195	165	105	100	45	31.47	27.84	35.35	27.19	23.75	30.93	17.61	14.74	20.91	16.50	13.70	19.73	7.23	5.42	9.57
	Q2	120	180	135	110	80	19.01	16.09	22.32	28.97	25.50	32.70	22.14	19.01	25.62	17.72	14.88	20.98	12.16	9.80	15.00

	Q3	100	155	130	115	90	16.36	13.56	19.60	26.86	23.38	30.65	22.59	19.35	26.19	19.90	16.84	23.36	14.30	11.70	17.36
	Q4	85	105	115	160	155	13.59	11.09	16.55	16.99	14.20	20.20	19.48	16.50	22.84	26.56	23.18	30.23	23.38	20.16	26.93
	Q5	45	90	110	135	230	7.42	5.58	9.81	14.86	12.21	17.96	18.32	15.40	21.66	22.63	19.44	26.18	36.76	32.94	40.76
	Survived 5 years (N=1,755)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	100	95	50	50	25	30.36	25.47	35.74	29.57	24.74	34.91	16.39	12.65	20.98	15.79	12.11	20.32	7.89	5.39	11.42
	Q2	55	95	75	65	55	15.71	12.24	19.95	27.67	23.17	32.67	22.64	18.50	27.40	18.49	14.70	23.00	15.49	12.02	19.74
	Q3	55	90	75	75	65	14.97	11.60	19.11	25.22	20.89	30.09	21.58	17.54	26.25	21.38	17.36	26.03	16.86	13.29	21.16
	Q4	50	50	70	100	95	13.75	10.56	17.72	14.20	10.95	18.22	19.43	15.64	23.88	27.19	22.82	32.05	25.43	21.16	30.23
	Q5	25	45	60	80	155	6.14	4.09	9.11	12.36	9.32	16.21	16.67	13.14	20.93	22.99	18.91	27.65	41.84	36.78	47.08
	Died within 5 years (N=1,310)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	95	70	55	50	20	33.11	27.84	38.85	24.40	19.74	29.77	18.86	14.70	23.86	17.08	13.11	21.94	6.55	4.22	10.01
	Q2	65	85	55	45	25	23.27	18.60	28.70	30.61	25.38	36.39	21.39	16.88	26.71	16.65	12.64	21.61	8.09	5.39	11.96
	Q3	45	65	55	40	30	17.99	13.55	23.48	29.06	23.54	35.28	24.02	18.94	29.97	17.85	13.43	23.33	11.08	7.68	15.73
	Q4	35	55	50	65	55	13.28	9.62	18.05	20.86	16.25	26.37	19.38	14.93	24.78	25.49	20.46	31.28	20.98	16.36	26.49
	Q5	25	45	50	55	75	9.40	6.30	13.80	18.21	13.82	23.62	20.66	15.97	26.29	21.91	17.11	27.60	29.83	24.32	35.99
Atlantic provinces (Newfoundland, PEI, Nova Scotia, New Brunswick)																					
Neighbourhood income quintiles							**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**
	Overall (N=2,350)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	155	120	95	65	55	30.81	26.81	35.11	23.93	20.31	27.97	20.40	17.01	24.27	13.46	10.67	16.83	11.40	8.85	14.58
	Q2	100	120	105	95	65	20.69	17.27	24.59	25.09	21.38	29.20	22.44	18.91	26.42	19.39	16.08	23.20	12.38	9.74	15.62
	Q3	95	120	110	90	80	18.41	15.21	22.12	24.01	20.39	28.05	22.77	19.22	26.76	19.18	15.89	22.98	15.62	12.64	19.15
	Q4	80	85	75	95	100	17.74	14.39	21.66	19.64	16.13	23.69	17.25	13.95	21.16	22.35	18.64	26.57	23.02	19.25	27.27
	Q5	60	80	85	85	135	13.46	10.53	17.05	17.90	14.54	21.83	19.69	16.20	23.73	19.29	15.83	23.29	29.66	25.50	34.20
	Survived 5 years (N=1,315)																				

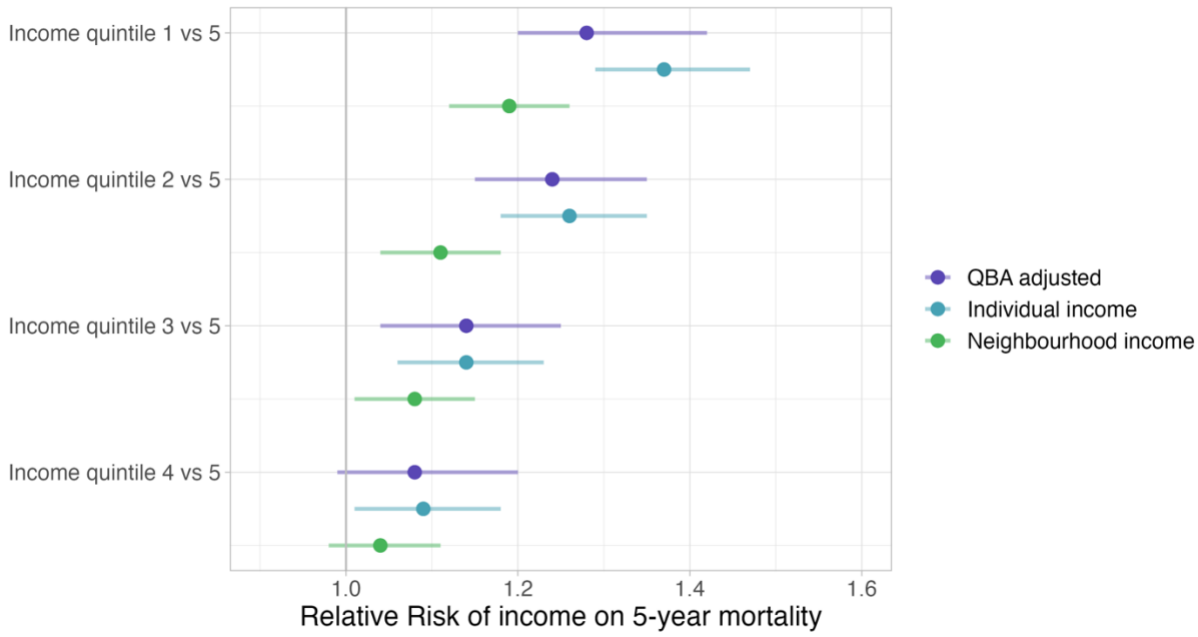
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	75	65	50	35	35	28.54	23.26	34.47	24.51	19.56	30.24	19.43	14.96	24.85	14.42	10.57	19.37	13.11	9.44	17.92
	Q2	50	65	65	55	35	17.63	13.49	22.71	24.59	19.76	30.16	24.22	19.42	29.77	20.41	15.96	25.72	13.14	9.59	17.75
	Q3	45	65	65	45	45	15.77	11.86	20.65	23.32	18.58	28.85	25.15	20.23	30.80	18.23	13.98	23.43	17.53	13.37	22.64
	Q4	35	50	50	60	75	12.73	9.22	17.32	18.09	13.88	23.24	19.09	14.77	24.31	23.07	18.36	28.56	27.03	21.98	32.74
	Q5	30	35	45	45	95	11.84	8.31	16.59	13.70	9.91	18.63	18.45	14.08	23.82	17.70	13.42	22.98	38.31	32.38	44.61
	Died within 5 years (N=1,040)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	80	55	50	30	25	33.46	27.62	39.86	23.16	18.14	29.08	21.42	16.57	27.22	12.23	8.59	17.13	9.72	6.51	14.27
	Q2	50	55	45	40	25	24.39	19.02	30.69	25.61	20.15	31.95	20.34	15.43	26.31	17.86	13.27	23.60	11.81	8.14	16.83
	Q3	50	55	45	45	30	21.71	16.74	27.66	24.74	19.45	30.92	19.91	15.12	25.76	19.99	15.19	25.83	13.66	9.72	18.86
	Q4	45	35	25	35	30	25.42	19.34	32.63	21.74	16.10	28.68	14.16	9.59	20.40	21.27	15.63	28.25	17.42	12.33	24.04
	Q5	30	45	40	40	35	15.64	11.08	21.60	23.47	17.90	30.13	21.20	15.91	27.66	21.03	15.76	27.50	18.67	13.67	24.96
Prairie provinces (Manitoba, Saskatchewan, Alberta)																					
Neighbourhood income quintiles							**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**	**Predictive value**	**Lower 95%**	**Upper 95%**
	Overall (N=3,985)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	290	185	150	90	60	37.18	33.76	40.73	23.65	20.75	26.83	20.02	17.30	23.06	11.58	9.48	14.07	7.56	5.90	9.64
	Q2	185	210	180	155	100	21.89	19.16	24.89	25.43	22.53	28.57	22.45	19.69	25.47	18.68	16.12	21.53	11.55	9.54	13.93
	Q3	165	185	190	155	135	19.36	16.77	22.24	22.24	19.49	25.27	23.68	20.85	26.75	18.86	16.29	21.72	15.87	13.50	18.56
	Q4	100	145	180	160	155	13.71	11.38	16.42	19.91	17.14	23.01	25.37	22.32	28.69	21.09	18.27	24.22	19.91	17.17	22.98
	Q5	100	130	150	165	265	11.88	9.81	14.31	15.81	13.43	18.52	19.51	16.88	22.44	20.53	17.85	23.50	32.26	29.06	35.64
	Survived 5 years (N=2,330)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	130	100	80	55	40	32.71	28.27	37.48	23.98	20.04	28.41	20.66	16.96	24.94	13.53	10.51	17.25	9.12	6.69	12.31
	Q2	85	105	110	105	70	18.14	14.90	21.91	21.84	18.31	25.83	23.56	19.92	27.64	22.55	18.97	26.58	13.91	11.07	17.34
	Q3	90	85	115	95	90	18.79	15.50	22.60	17.47	14.29	21.19	24.81	21.09	28.95	20.36	16.93	24.28	18.56	15.26	22.38

	Q4	50	75	115	110	110	10.58	8.05	13.78	16.12	13.00	19.83	26.02	22.17	30.28	23.82	20.11	27.98	23.45	19.77	27.59
	Q5	50	75	100	110	180	10.14	7.79	13.10	14.07	11.30	17.38	20.01	16.73	23.76	21.59	18.20	25.41	34.19	30.14	38.48
	Died within 5 years (N=1,655)																				
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**			**Q2**			**Q3**			**Q4**			**Q5**		
	Q1	155	85	65	35	25	42.84	37.68	48.17	23.05	18.93	27.75	19.05	15.27	23.50	9.28	6.68	12.74	5.79	3.83	8.65
	Q2	100	105	70	50	30	26.96	22.48	31.96	30.36	25.68	35.48	20.77	16.78	25.43	13.51	10.27	17.56	8.41	5.92	11.81
	Q3	75	95	75	60	45	20.63	16.63	25.30	28.40	23.82	33.48	21.83	17.74	26.57	16.75	13.15	21.10	12.38	9.32	16.28
	Q4	55	70	65	50	45	18.79	14.56	23.90	25.87	20.98	31.43	23.86	19.15	29.31	16.93	12.94	21.85	14.55	10.88	19.21
	Q5	50	55	50	55	90	14.76	11.16	19.26	18.58	14.52	23.48	18.13	14.09	23.02	19.01	14.86	23.99	29.52	24.48	35.12

Supplemental Table 4. Bias adjusted relative risks without random error and 95% simulation intervals for all of Canada and by province
* Using beta distribution created from confidence intervals of the adjusted predicted values
* 10,000 simulations

	Differential
Scenario	**Simulation RR without random error (95% SI)**
Canada	
Quintile 1	1.42 (1.33-1.52)
Quintile 2	1.35 (1.26-1.44)
Quintile 3	1.17 (1.09-1.26)
Quintile 4	1.11 (1.03-1.19)
Ontario	
Quintile 1	1.40 (1.28-1.53)
Quintile 2	1.32 (1.16-1.44)
Quintile 3	1.11 (1.00-1.22)
Quintile 4	1.07 (0.96-1.19)
Quebec	
Quintile 1	1.36 (1.18-1.68)
Quintile 2	1.30 (1.13-1.57)
Quintile 3	1.24 (1.07-1.48)
Quintile 4	1.12 (0.94-1.35)
British Colombia	
Quintile 1	1.40 (1.17-1.68)
Quintile 2	1.34 (1.13-1.60)
Quintile 3	1.29 (1.08-1.56)
Quintile 4	1.19 (0.99-1.45)
Prairie provinces	
Quintile 1	1.52 (1.30-1.79)
Quintile 2	1.48 (1.27-1.75)
Quintile 3	1.20 (1.01-1.43)
Quintile 4	1.06 (0.87-1.28)
Atlantic provinces	
Quintile 1	1.50 (1.23-1.85)
Quintile 2	1.37 (1.12-1.69)
Quintile 3	1.23 (0.99-1.53)
Quintile 4	1.26 (1.01-1.58)

Supplemental Figure 1. Applying the Canada-wide bias parameters to the 2001 cohort and comparing bias-adjusted relative risks to the relative risk of household and neighbourhood income on 5-year mortality for all of Canada



Supplemental Table 5. Cohort characteristics compared to all colorectal cancer patients diagnosed from 2006 to 2014 in the Canadian Cancer Registry (only compared on variables found in the CCR)

	CCR diagnosed 2006-2014 (N=165,155)	**2006 Census (21,595)**	**Standardized mean difference**
Variabes			
Age at diagnosis (mean (SD))	69.80 (12.62)	65.95 (12.46)	0.307
diagnosis year			
2006	19240 (11.7)	2620 (12.1)	0.099
2007	19930 (12.1)	2750 (12.7)	
2008	20505 (12.4)	2960 (13.7)	
2009	20670 (12.5)	2890 (13.4)	
2010	20835 (12.6)	3000 (13.9)	
2011	15755 (9.5)	1800 (8.3)	
2012	15860 (9.6)	1870 (8.7)	
2013	16060 (9.7)	1895 (8.8)	
2014	16300 (9.9)	1810 (8.4)	
Sex			
male	90655 (54.9)	11945 (55.3)	0.009
female	74505 (45.1)	9650 (44.7)	
Tumour location			
rectal	55090 (33.4)	7245 (33.6)	0.004
colon	110070 (66.6)	14350 (66.4)	
Stage at diagnosis			
0-1	16905 (10.2)	2115 (9.8)	0.064
2	18915 (11.5)	2400 (11.1)	
3	21375 (12.9)	2580 (11.9)	
4	15150 (9.2)	1790 (8.3)	
5	4030 (2.4)	460 (2.1)	
Missing/unknown	88785 (53.8)	12255 (56.7)	
Province/territory at diagnosis			
newfoundland and Labrador	4575 (2.8)	650 (3.0)	0.043
PEI	955 (0.6)	125 (0.6)	
Nova scotia	6885 (4.2)	945 (4.4)	
New Brunswick	4790 (2.9)	630 (2.9)	
Quebec	26370 (16.0)	3420 (15.8)	
Ontario	67885 (41.1)	8750 (40.5)	
Manitoba	7160 (4.3)	1020 (4.7)	

Saskatchewan	6275 (3.8)	910 (4.2)	
Alberta	15445 (9.4)	2055 (9.5)	
British Columbia	24375 (14.8)	3055 (14.1)	
Territories	445 (0.3)	35 (0.2)	

Chapter 5 : Methods for defining the diagnostic interval

5.1 Preface

This chapter moves away from the measurement of income to describe the methodology I used to define the diagnostic interval in more detail than could be provided in the manuscript. In this chapter, using a cohort of colon cancer patients, I expand on an existing method by Groome et al. Groome and colleagues to define the diagnostic interval using administrative data. This approach adopts variable lookback periods for each type of patient encounter based on statistical process control, offering a more nuanced, data-driven method to define diagnostic intervals.

5.2 Diagnostic interval methods

I performed this method in a retrospective cohort of colon cancer patients (ICD-O-3 C18.0, C18.2-C18.9) diagnosed in Ontario, Canada, between January 1, 2007 and December 31, 2019. I excluded individuals who had a death date before their diagnosis date, those diagnosed with multiple cancers on the same day, and those with less than two years of Ontario Health Insurance Plan (OHIP) eligibility before diagnosis. I used linked administrative databases obtained from ICES (formerly the Institute for Clinical Evaluation Sciences), which are described in detail in Table 5.1. The cohort consisted of 65,049 colon cancer patients diagnosed from 2007 to 2019. I could not identify an index contact date in 504 (0.8%) patients and excluded an additional 220 patients for whom demographic variables were missing for a final cohort of 64,303. The mean age at diagnosis was 71 years (SD 13) and 48.8% were female. Cancer stage distribution was 18.9% stage I, 25.0% stage II, 24.1% stage III, 17.4% stage IV and 14.6% unknown stage.

TABLE 5.1 ADMINISTRATIVE DATA SOURCES USED TO CREATE THE DIAGNOSTIC INTERVAL

Database	Description
Ontario Cancer Registry	Record for all primary cancers in Ontario residents. Used to create the cohort and for the diagnosis date to calculate the end of the diagnostic interval. Includes cancer site, diagnosis date and cancer stage.
Ontario Health Insurance Plan Database	Billing claims made by all Ontario physicians, including inpatient and outpatient settings. Each claim includes the date, one fee code representing the billable service and one diagnosis code, physician specialty and referring physician where applicable. Physicians are required to submit a diagnosis with each fee code.
CIHI Discharge Abstract Database	Hospital discharge abstracts for Ontario, including administrative, demographic and clinical data. Each record includes up to 20 intervention codes and 25 diagnosis codes.

CIHI Same Day Surgery Database	All same-day surgery or procedure stays in all Ontario day surgery clinics and institutions. Includes administrative, demographic, and clinical data. Each record includes up to 20 intervention codes and 25 diagnosis codes.
CIHI National Ambulatory Care Reporting System	All emergency department visits in Ontario. Includes administrative, demographic, and clinical data. Each record includes up to 10 intervention codes and 10 diagnosis codes.
Registered Persons Database	Provides basic demographic information. Used to measure age, sex, rural residence and neighbourhood income quintile.

I used a data-driven approach to identify the earliest cancer-related encounter and calculate the diagnostic interval in a cohort of patients already diagnosed with colon cancer. Encounters were defined as unique patient contacts with the healthcare system (e.g., physician visits, hospital admissions) and were identified from relevant data sources using diagnosis and procedure codes (OHIP schedule of benefits, ICD-9, ICD-10 and CCI codes). The goal was to identify the first encounter that could be reasonably attributed to the colorectal cancer diagnosis. To do this, I identified encounters that occurred more frequently in the 0-3 months compared to the 24-27 months before diagnosis and used statistical process control to identify specific lookback periods for each group of similar encounters. I collected all encounters for each patient that occurred during the lookback periods specific to each encounter category and identified the earliest encounter as the first contact date. If that earliest encounter was a procedure, I looked back a further 365 days for a referring physician and identified the closest referral date as the first contact date. The end of the interval was defined as the diagnosis date in the Ontario Cancer Registry (OCR), which is normally the first positive biopsy date but in the absence of pathology takes dates from hospitalization and death certificate data.[1] I checked the diagnostic interval results by symptomatic and asymptomatic diagnostic pathways to determine validity.

Asymptomatic pathways were defined as having a guaiac fecal occult blood test (gFOBT) or lower gastrointestinal (GI) scope on the first contact date without another procedure and without a visit to the emergency department (ED) or with an emergency physician on that date.[2,3] Symptomatic pathways included all other pathways.

These methods are described in more detail below and differences between my approach and the Groome approach are detailed in Supplemental Table 1.

Step one: Identify relevant physician specialties in physician billing data.

First, I determined physician specialties that could be involved with a colon cancer diagnostic work up by identifying the specialties seen more often in the 0-3 months compared to the 24-27 months before diagnosis. Physician specialty was self-reported in the OHIP database. I limited the physician visit and procedure dataset to only those encounters occurring with specialties that were seen by at least 1% of the cohort in the 0-3 months before diagnosis and had at least a 20% increase between the two periods. The physician billing database was limited to diagnosis and procedure codes billed by 22 physician specialties, listed in Table 5.2.

TABLE 5.2 PHYSICIAN SPECIALTIES SEEN BY AT LEAST 1% OF THE COHORT IN MONTHS 0-3 AND THAT SHOWED AT LEAST A 20% INCREASED BETWEEN DIAGNOSIS AND CONTROL PERIODS*

	0-3 months before diagnosis		24-27 months before diagnosis		
Physician specialty, as assigned by OHIP for payment purposes	**Count (N)**	**Proportion (%)**	**Count (N)**	**Proportion (%)**	**Percent change**
Family practice and general practice	71407	93.12%	50139	65.39%	42.42%
Diagnostic radiology	53906	70.30%	17274	22.53%	212.06%
General surgery	46399	60.51%	3103	4.05%	1395.29%
Internal medicine	39596	51.64%	11797	15.38%	235.64%

Anaestheia	24313	31.71%	2465	3.21%	886.33%
Gastroenterology	18981	24.75%	877	1.14%	2064.31%
Cardiology	15222	19.85%	5345	6.97%	184.79%
Emergency medicine	6683	8.72%	941	1.23%	610.20%
Urology	4615	6.02%	2958	3.86%	56.02%
Pathology, microbiology, clinical biochemistry	3576	4.66%	1519	1.98%	135.42%
Respiratory disease	2343	3.06%	1075	1.40%	117.95%
Obstetrics and gynaecology	2110	2.75%	1231	1.61%	71.41%
Haematology	1845	2.41%	655	0.85%	181.68%
Neurology	1662	2.17%	1240	1.62%	34.03%
Nuclear medicine	1648	2.15%	703	0.92%	134.42%
Nephrology	1057	1.38%	494	0.64%	113.97%
Therapeutic radiology	1035	1.35%	706	0.92%	46.60%
Medical oncology	1018	1.33%	239	0.31%	325.94%
Rheumatology	874	1.14%	688	0.90%	27.03%
Geriatrics	867	1.13%	338	0.44%	156.51%
Thoracic surgery	857	1.12%	146	0.19%	486.99%
Endocrinology	826	1.08%	515	0.67%	60.39%

*Excluded physician specialties were ophthalmology, dermatology, orthopaedic surgery, otolaryngology, psychiatry, plastic surgery, cardiothoracic surgery, vascular surgery, physical medicine, infectious disease, microbiology, neurosurgery, pediatrics, nurse practitioners, non-medical professionals, and clinical immunology.

Step two. Identify and categorize relevant encounters.

Using the limited physician billing data from step one, I identified and categorized healthcare encounters that could be attributed to the diagnosis of colon cancer. First, I compared the number of encounters that occurred in the diagnosis period (0-3 months before diagnosis) and control periods (24-27 months before diagnosis). These periods were hypothesized as the time where encounters would be most and least likely related to colon cancer, respectively. I calculated the percent change in encounter frequency: the ratio of the difference in encounters between the control and the diagnosis period multiplied by 100. I then limited the dataset to only encounters with diagnosis and procedure codes that were assigned to at least 100 patients in the

diagnosis period and had ≥20% increase in frequency between the control and diagnosis periods. I further excluded codes that did not represent patient encounters, this included: add-on codes, encounters where the patient was not present (e.g., tracking codes, physician to physician consultations), premiums (e.g., weekend fees), subsequent encounters (e.g., second consultation) and concurrent encounters (e.g., guidance of a biopsy).

Next, I categorized diagnosis and procedure codes that were similar to each other to create encounter categories (for example, I combined general surgery consultation and assessment codes into one category "General surgery consultation"), which were used to create lookback periods in each category. I used categories previously identified by Webber et al[3,4], which were created with input from a gastroenterologist and a family physician, and created new categories based on the Ontario Ministry of Health diagnoses and input from a colorectal surgeon (SP).[5]

Using the diagnosis and procedure codes from OHIP, I created a cross-walk to identify similar ICD-10 diagnosis and Canadian Classification of Health Interventions (CCI) procedure codes in the CIHI hospitalization databases (NACRS, DAD and SDS). I then created a new dataset that contained all eligible encounters in OHIP, NACRS, DAD and SDS occurring in the 0-4 years before diagnosis. If an individual had multiple encounters on the same day with the same encounter category, only the OHIP encounter was used. I prioritized OHIP because this dataset contains information on the referring physician which was used to identify the first contact date. In this dataset, I created new diagnosis and control periods of 0-3 years and 3-4 years respectively, to allow for longer lookback periods where appropriate.

The diagnosis and procedure codes that demonstrated a 20% increased between control and diagnosis periods and their categorizations are listed in Supplemental Tables 2a and 2b. I

generated 13 diagnosis and 26 procedure categories. Diagnosis encounter categories included colorectal cancer and other cancer categories and signs and symptoms. A diagnosis for colorectal or other cancer can indicate either confirmed diagnosis or suspicion of cancer based on the patients' symptoms. Procedure encounter categories included consultations, imaging procedures, surgical procedures, and other procedures.

Step three: Use statistical process control to derive encounter category-specific lookback periods.

Using the dataset of encounters created in step two, I determined unique lookback periods for each diagnosis and procedure category using statistical process control methods.[6–8] Statistical process control identifies variation that lies outside of an expected average rate, for example, when the weekly colonoscopy rate is higher than the average weekly rate in the control period.[6] This method allowed us to define a non-arbitrary number of weeks for which to look for each encounter category prior to diagnosis. I defined lookback periods across three time frames (2007-2011, 2012-2015, 2016-2019) to account for the evolution in diagnostic approaches across time. An additional timeframe was considered only for guaiac fecal occult blood test (gFOBT) to align with changes to population-level colon cancer screening delivery (2007-2008, 2009-2011, 2012-2015, 2016-2019).[9,10]

Using statistical process control, for each encounter category in each time period, I compared the weekly frequency of encounters in the diagnosis period (0-3 years before diagnosis) to the background encounter rate. The background encounter rate was defined as the mean weekly encounter frequency in the control period (3-4 years before diagnosis). I then applied the control chart rules at each week in the diagnosis period to determine at which week

the weekly count would exceed the background rate. The rules are detailed below, with each rule becoming progressively more liberal.

Rule 1. A weekly count equal to or greater than three standard deviations greater than the background encounter rate. The probability of this happening by chance is rare, for example, if I assume a normal distribution, there is a 0.03% chance of finding a value beyond 3 standard deviations.

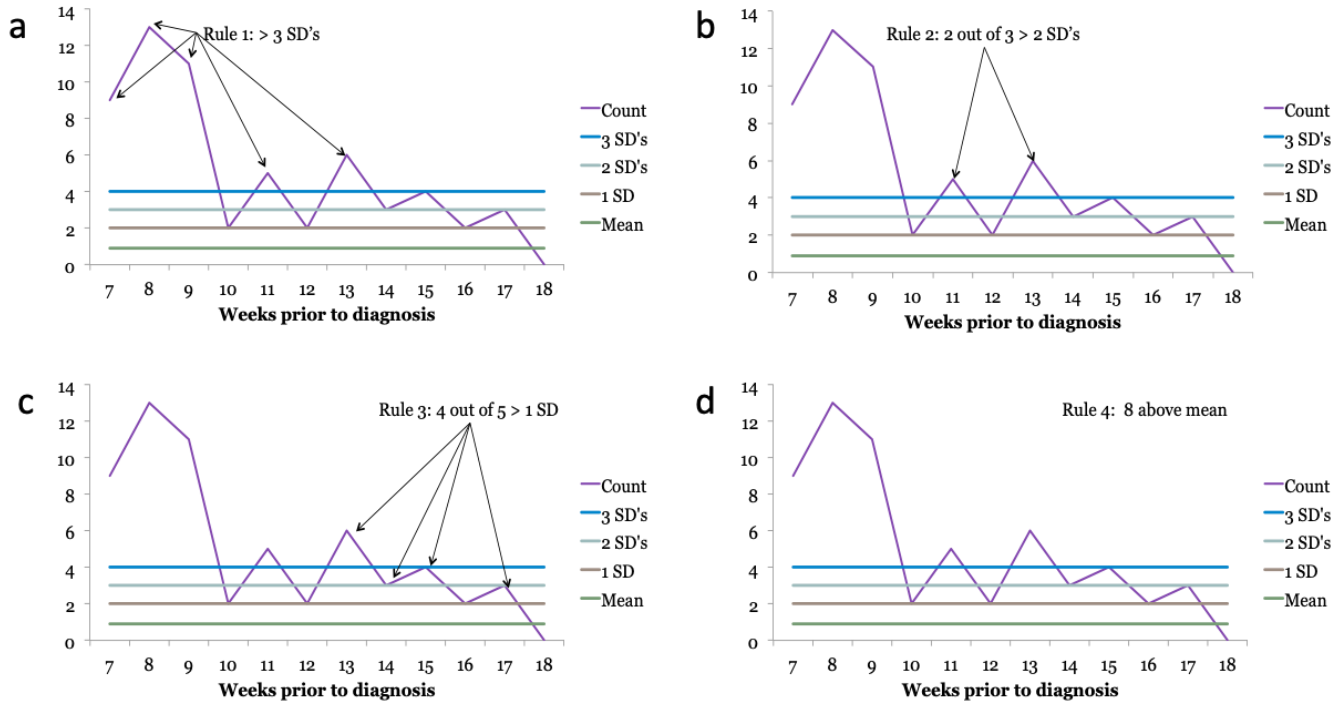
Rule 2. Two out of three successive weekly counts greater than two standard deviations more than the background encounter rate.

Rule 3. Four out of five successive weekly counts greater than one standard deviation more than the background encounter rate.

Rule 4. Eight successive values greater than the background encounter rate.

I calculated four lookback periods for each encounter category using rules 1 to 4, rules 1 to 3, rules 1 to 2 and rule 1 only, with each providing an increasingly shorter lookback period. The consecutive week at which none of the rules could be applied marked the end of the lookback period for that encounter category, even if the rules applied to a non-consecutive week further back in time. A visual example of how to apply the control chart rules can be found in Figure 5.1.

FIGURE 5.1 EXAMPLE OF HOW CONTROL CHART RULES ARE APPLIED (A=RULE 1; B=RULE 2; C=RULE 3; D=RULE 4



*Figures obtained from: Groome P, Whitehead M, Grunfeld E, Moineddin R, Irish J. Initial Peri-Diagnostic Encounter Leading to Cancer Diagnosis, development of an administrative data-based approach. ICES Clinical Round Presentation. May 2014.

Step four: Determine each encounter's signal strength.

I used the equation below to calculate the signal strength for each of the four combinations of rules for each encounter category. The signal strength is the proportion of encounters in the lookback period that exceeded the expected number of encounters based on the background encounter rate.

$$C = total\ number\ of\ encounters\ from\ x\ weeks\ before\ diagnosis$$
$$B = mean\ background\ rate\ in\ the\ control\ period * x\ weeks$$
$$A = C - B$$
$$Signal\ strength = \left(\frac{A}{C}\right) * 10$$

I set a cut-off of 70% to consider signal strength as adequate and encounter categories that did not demonstrate a signal strength above 70% for any of the four rule combinations were

excluded in subsequent steps and did not contribute to the calculation of the diagnostic interval length. gFOBT was the exception given its role in the colon cancer diagnostic process and was included regardless of signal strength. I applied the rules in a stepwise process, beginning with the most liberal set of rules (rules 1 to 4) and ending with the least liberal (rule 1). The first week where the set of rules reached 70% was used as the lookback period in the definition of the diagnostic interval. For example, if rules 1 to 4 reached 65%, rules 1 to 3 reached 70% and rules 1 to 2 reached 80%, the week at which rules 1 to 3 applied would be the lookback period. I provide an example in Table 5.3.

TABLE 5.3 EXAMPLE OF DATA STRUCTURE AND CONTROL CHART RULES

Signal strength for rule 1-3 calculated by:

$C = total\ number\ of\ encounters\ from\ x\ weeks\ before\ diagnosis$

$B = mean\ background\ rate\ in\ the\ control\ period * x\ weeks$

$A = C - B$

$$Signal\ strength = \left(\frac{A}{C}\right) * 100$$

A = 22358 – 421.42 = 21936.58

B = 11.09 * 38 weeks = 421.42

C = 22358

Signal strength = (21936.58/22358) * 100 = 98.11%

Observation	Week	Group	Count	Standardized count	Cumulative count	Background rate*	Rule	Signal strength
1	1	1	16557	4492.41	16557	11.09	1	
2	2	1	2098	566.62	18655	11.09	1	
3	3	1	902	241.89	19557	11.09	1	
4	4	1	525	139.53	20082	11.09	1	
5	5	1	327	85.77	20409	11.09	1	
…								
34	34	1	27	4.32	22288	11.09	1	
35	35	1	26	4.05	22314	11.09	1	98%
36	36	1	17	1.60	22331	11.09	2	98%
37	37	1	9	-.57	22340	11.09	3	
38	38	1	18	1.87	22358	11.09	3	98%
39	39	1	11	-.03	22369	11.09	.	
40	40	1	11	-.03	22380	11.09	.	
…								
153	153	1	8	-.84	24005	11.09	.	

154	154	1	10	-.30	24015	11.09	.	
155	155	1	21	2.69	24036	11.09	.	
156	156	1	10	-.30	24046	11.09	.	
157	157	1	8	-.84	24054	11.09	.	

*Background rate = Mean weekly count in control period

Final lookback periods and signal strengths for each encounter category are described in Table 5.4. Ten encounter categories were discarded because they did not reach 70% signal strength. Lookback periods for each encounter category differed by the timeframe of diagnosis, for example, the category for abdominal CT had a lookback period of 37 weeks between 2007 and 2011, 49 weeks between 2012 and 2015 and 89 weeks between 2016 and 2019. Of the 29 encounter categories, 21 (72%) had lookback periods for all three time periods occurring within one year. The diagnosis encounter category for anemia was the only category that had a lookback period longer than two years (115 weeks in 2007-2011 and 112 weeks in 2016-2019). The mean weekly encounter frequency in the 3 years before diagnosis was highest for non-emergency family physician visits, signs and symptoms not otherwise specified (NOS) and signs and symptom GI and lowest for abdominal or pelvis MRI and signs and symptoms liver.

TABLE 5.4 LOOKBACK PERIODS AND SIGNAL STRENGTH FOR EACH ENCOUNTER CATEGORY. GREEN HIGHLIGHTING INDICATES THE CATEGORY HAD 70% SIGNAL STRENGTH AND WAS INCLUDED IN THE DIAGNOSTIC INTERVAL DEFINITION

		2007-2011/2012-2015/2016-2019			
	Category	**Lookback period (weeks)**	**gnal strength (%)**	**Mean weekly encounter frequency in the 0-36 months before diagnosis**	**Mean weekly encounter frequency in the 36-47 months before diagnosis**
	Diagnosis groups				
1	Colorectal cancer	49/37/50	98/97/98	64/46/43	4/5/2
2	Other cancer	37/24/25	77/79/78	69/60/69	37/36/40
3	Signs and symptoms GI	64/50/62	71/75/72	336/253/242	169/129/118
4	Signs and symptoms liver	26/19/32	74/82/71	7/8/9	5/5/5
5	Signs and symptoms Haematology	36/39/62	78/77/71	51/44/41	25/21/20
6	Anemia	115/89/112	78/83/80	63/51/50	17/13/13
7	Signs and symptoms nutritional	7/6/5	70/71/73	22/19/19	20/16/15
8	Signs and symptoms NOS	14/22/24	61/55/53	519/443/446	420/344/348
9	Bacterial or viral	13/13/9	79/79/83	10/9/11	7/6/7
10	Neurological disorders	8/7/9	52/55/50	156/118/107	135/98/89
11	Circulatory system disorders or symptoms	7/7/10	58/62/57	299/231/203	251/187/165
12	Genitourinary system disorders or symptoms	7/9/9	52/50/51	93/78/75	79/65/63
13	Respiratory system disorders or symptoms	8/9/10	54/59/57	99/77/77	84/63/60
	Procedure groups				
17	Abdominal or pelvis MRI	11/13/11	96/95/96	1/2/2	.35/.54/1
18	Biopsy	48/42/39	97/97/97	63/60/66	6/6/6
19	Abdominal ultrasound	11/11/10	73/72/71	52/46/47	38/34/36
20	Other ultrasound	1/2/2	47/58/61	9/9/16	8/9/12
21	Head MRI	4/3/3	79/75/71	9/11/12	6/9/12
22	Non-emergency family physician visit	12/11/13	31/31/33	1675/1172/1056	1613/1140/987
23	Emergency family physician visit	25/25/26	73/73/70	209/191/188	137/118/125
24	Critical care	47/74/65	88/83/79	18/17/19	5/5/7
25	Internal medicine consult	9/10/13	70/70/67	145/88/84	124/76/64
26	Gastroenterologist consult	61/47/38	76/76/78	21/18/20	9/9/10
27	Lower GI scope	37/35/36	95/95/95	133/107/104	24/21/19
28	Upper GI scope	49/40/39	93/93/94	51/50/60	10/12/13
29	Non-GI scope	4/4/5	79/80/80	21/18/16	18/14/13
30	gFOBT*	47/53/25/24	80/67/65/73	34/69/84/76	14/35/68/54

31	Colon resection	31/18/21	99/100/100	50/35/31	2/1/1
32	Cancer visit	1/6/6	50/68/69	16/21/27	13/16/20
33	Abdominal Xray	14/12/14	83/84/80	210/169/159	128/104/99
34	Other abdominal procedure	22/18/26	94/96/93	15/12/11	4/3/3
35	Miscellaneous procedure	7/4/4	79/72/72	9/10/11	5/7/8
36	Lung procedure	4/3/5	64/61/67	14/11/13	11/10/10
37	Abdominal CT	37/49/89	93/89/83	82/76/85	19/20/22
38	Head CT	25/25/22	74/75/77	38/38/41	23/22/24
39	Other Xray	1/1/1	40/42/1	25/31/38	24/24/38
40	General surgery consultation	35/38/47	77/74/71	92/65/60	52/38/34
41	Other consultation	11/10/1	73/74/66	172/161/181	134/119/129
42	Cardiovascular visit	9/9/7	76/73/75	190/174/182	145/132/141

*gFOBT time periods are 2007-2008, 2009-2011, 2012-2015, 2016-2019

Step five: Collect all eligible cancer-related encounters before diagnosis using specific lookback periods for each encounter category.

For each patient in the cohort, I captured all eligible encounters and their dates that occurred within the unique lookback period for each encounter category in each time period.

Step six: Add referring physician encounter date to each procedure category.

To capture the full diagnostic interval, I identified the referring physicians for all physician procedure-based encounters that were identified in the billing data. I looked back up to 365 days from the date of the procedure and chose the closest encounter date with that referring physician.

Step seven: Determine the first contact date and calculate the diagnostic interval.

I defined the first contact date by ordering all eligible and identified encounter categories per person from closest to diagnosis to furthest and defined the earliest encounter as the first contact date. If the first encounter was a procedure with a referring physician, then the referring physician date was defined as the first contact date. If a referring physician visit could not be identified, the procedure date was defined as the first contact date. I calculated the diagnostic interval as the number of days between the first contact date and cancer diagnosis date, with a minimum of 1 day.

Table 5.5 details the diagnostic interval distribution for the whole cohort and by diagnostic pathway symptom status. The median diagnostic interval overall was 108 days (IQR 31-243 days) with a 90th percentile of 383 days. 17.7% of patients had asymptomatic diagnostic pathways and asymptomatic pathways had shorter median and 90th percentile diagnostic intervals compared to symptomatic pathways (median 71 days (IQR 35-137, 90th percentile 230 days) vs

median 121 days (IQR 29-273, 90th percentile 404 days), respectively). Earlier stage at diagnosis had longer diagnostic intervals than later stage at diagnosis.

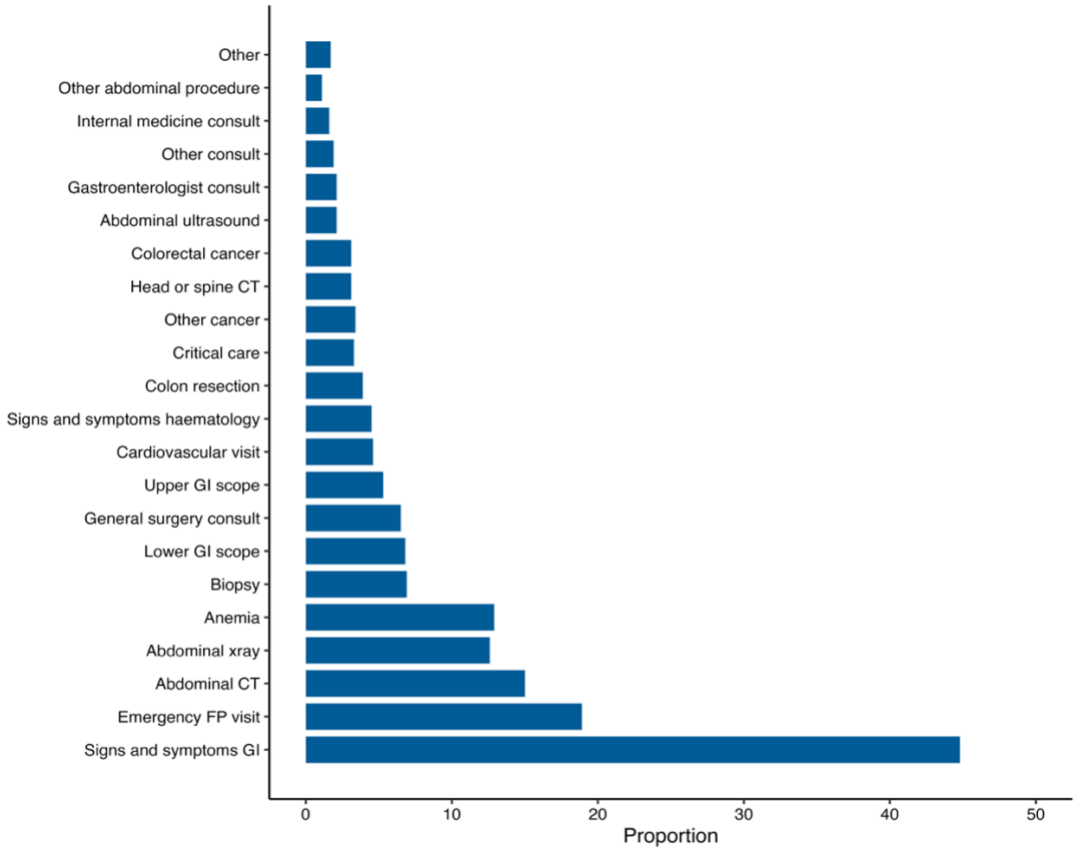
TABLE 5.5 MEDIAN AND 90TH PERCENTILE DIAGNOSTIC INTERVAL BY SYMPTOM STATUS, STAGE AT DIAGNOSIS AND DIAGNOSIS YEARS

	Overall (N=64,303)			Asymptomatic (N=11,378)			Symptomatic (N=52,925)		
Variable	**N**	**Median (IQR)**	**90th percentile (95% CI)**	**N**	**Median (IQR)**	**90th percentile (95% CI)**	**N**	**Median (IQR)**	**90th percentile (95% CI)**
Diagnostic interval overall	64,303	108 (31-243)	383	11,378	71 (35-137)	230	52,925	121 (29-273)	404
Stage at diagnosis									
Stage I	12,126	128 (51-252)	375	3,071	82 (39-144)	224	9,055	154 (60-287)	405
Stage II	16,062	107 (33-238)	381	2,861	66 (35-133)	223	13,201	120 (32-267)	401
Stage III	15,513	104 (31-239)	379	2,773	71 (35-134)	235	12,740	117 (29-268)	400
Stage IV	11,193	76 (15-210)	366	1,322	77 (33-153)	256	9,871	75 (13-219)	379
Stage unknown/missing	9,409	122 (34-286)	418	1,351	60 (29-124)	207	8,058	140 (36-309)	432
Diagnosis year									
2007	4,934	123 (35-265)	406	894	89 (39-182)	261	4,040	135 (34-294)	423
2008	5,025	113 (32-251)	391	1,106	84 (37-162)	244	3,919	128 (30-286)	410
2009	5,016	124 (37-269)	393	1,212	115 (52-222)	309	3,804	131 (33-291)	415
2010	4,946	119 (36-261)	395	1,029	97 (44-202)	306	3,917	127 (32-285)	415
2011	5,059	116 (37-255)	388	1,115	95 (45-196)	307	3,944	123 (34-279)	411
2012	4,943	99 (29-202)	314	805	65 (31-110)	158	4,138	110 (29-226)	323
2013	4,922	97 (28-201)	313	774	59 (30-107)	155	4,148	109 (28-224)	324
2014	4,940	93 (28-202)	316	799	59 (29-109)	157	4,141	104 (27-225)	331
2015	5,001	98 (29-208)	317	798	64 (33-115)	162	4,203	111 (28-234)	327
2016	4,957	105 (28-267)	411	739	56 (30-105)	152	4,218	123 (27-302)	424
2017	4,997	110 (29-278)	421	736	58 (29-105)	154	4,261	133 (30-310)	434
2018	4,837	110 (29-280)	418	651	56 (30-106)	149	4,186	131 (28-306)	432
2019	4,726	110 (31-278)	418	720	55 (32-97)	141	4,006	133 (31-310)	432

Step eight: Define the first encounter.

I defined the first encounter as that encounter occurring on the first contact date. If the first encounter was a referring physician, I defined the first encounter as the procedure preceding the referring physician visit. If a patient had multiple encounters on the first contact date, I described all encounters. The first encounter differed by symptomatic and asymptomatic pathways, with 76.1% of patients with asymptomatic pathways having a gFOBT on their index encounter date and the rest having a lower GI endoscopy. The most common index encounter for patients with a symptomatic pathway was an encounter with a diagnosis code for GI symptoms (44.5%), emergency family physician visit (19.3%) and abdominal CT (15.5%) (Figure 5.2).

FIGURE 5.2. FREQUENCY OF FIRST ENCOUNTER CATEGORY FOR PATIENTS WITH SYMPTOMATIC DIAGNOSTIC PATHWAYS (PATIENTS CAN HAVE MULTIPLE)



*****Other = signs and symptoms nutritional, signals and symptoms bacterial or viral, signs and symptoms liver, non-GI scope, abdominal MRI, gFOBT, miscellaneous procedure, head MRI

References

1. Clarke E, Marrett L, Kreiger N. Cancer registration in Ontario: A computer approach. In: Cancer registration principles and methods. Lyon, France: IARC Publication; 1991. p. 246–57.

2. Guan Z, Webber C, Flemming JA, Mavor ME, Whitehead M, Chen BE, et al. Real-world colorectal cancer diagnostic pathways in Ontario, Canada: A population-based study. Eur J Cancer Care (Engl). 2022;31(5):e13603.

3. Webber C. Availability and quality of colonoscopy resource and the colorectal cancer diagnostic interval [Internet] [PhD Thesis]. [Kingston, ON]: Queen's University; 2017. Available from: https://qspace.library.queensu.ca/handle/1974/6707

4. Webber C, Flemming J, Birtwhistle R, Rosenberg M, Groome P. Wait times and patterns of care in the colorectal cancer diagnostic interval: IJPDS (2017) Issue 1, Vol 1:191, Proceedings of the IPDLN Conference (August 2016). Int J Popul Data Sci [Internet]. 2017 Apr 18 [cited 2021 Aug 19];1(1). Available from: https://ijpds.org/article/view/211

5. Ontario Ministry of Health and Longterm Care. Diagnostic Codes [Internet]. Toronto, ON: Ontario Ministry of Health and Longterm Care; 2023 Apr. Available from: https://www.health.gov.on.ca/en/pro/programs/ohip/claims_submission/diagnostic_codes.pdf

6. Shewhart WA. Economic Control of Quality of Manufactured Product. D. Van Nostrand Company, Incorporated; 1931. 540 p.

7. Benneyan J, Lloyd R, Plsek P. Statistical process control as a tool for research and healthcare improvement. Qual Saf Health Care. 2003 Dec;12(6):458–64.

8. Mohammed M. Using statistical process control to improve the quality of health care. Qual Saf Health Care. 2004 Aug;13(4):243–5.

9. Sewitch MJ, Fournier C, Ciampi A, Dyachenko A. Adherence to colorectal cancer screening guidelines in Canada. BMC Gastroenterol. 2007 Oct 2;7:39.

10. Rabeneck L, Paszat LF. A Population-Based Estimate of the Extent of Colorectal Cancer Screening in Ontario. Off J Am Coll Gastroenterol ACG. 2004 Jun;99(6):1141.

11. Groome PA, Webber C, Whitehead M, Moineddin R, Grunfeld E, Eisen A, et al. Determining the Cancer Diagnostic Interval Using Administrative Health Care Data in a Breast Cancer Cohort. JCO Clin Cancer Inform. 2019 Dec 1;(3):1–10.

5.3 Supplemental material

Supplemental Table 1. Detailed steps and differences in creating the diagnostic interval between the Groome approach and my approach

Steps	Description	Original Groome/Webber et al[4,11]	Modified Davis et at
1. Identify relevant physician specialties	Identify physician specialties seen by patients more often in the 0-3 months immediately preceding diagnosis compared to a control period	• Control period: 18-21 and 21-24 months before diagnosis • Inclusion criteria: not reported	• Control period: 24-27 months before diagnosis • Inclusion criteria: 20% increase between 0-3 months and seen by at least 1% of the cohort in the 0-3 months before diagnosis
2. Identify and categorize encounters	Using the specialties identified in step one, identify OHIP procedure and diagnosis codes occurring more frequently in the 0-3 months before diagnosis compared to the control period.	• Control period: 18-24 months before diagnosis • Inclusion criteria: not reported • Kept only codes that were a priori assumed to be relevant to colorectal cancer	• Control period: 24-27 months before diagnosis • Inclusion criteria: 20% increase between 0-3 months and control period and occurs in at least 100 patients in the 0-3 months before diagnosis • Kept all codes that demonstrated an increase regardless of relevance to colon cancer
	Group together similar diagnoses and similar procedures.	• Created categories with input from a gastroenterologist and family physician • 9 diagnosis and 15 procedure categories	• Used previous categories and created new categories with input from a surgical oncologist and the Ontario Ministry of Health Diagnosis categories • 13 diagnosis and 26 procedure categories
	Identify diagnosis and procedure codes in CIHI databases (NACRS, DAD and SDS)	• Created a cross-walk from diagnosis and procedures in OHIP to identify similar codes in ICD-10 and CCI	• No differences
	Identify all encounter categories in OHIP	• 0-24 months before diagnosis	• 0-48 months before diagnosis

	and CIHI databases in the full lookback period		
3. Use statistical process control to derive category-specific encounter lookback periods	Calculate weekly counts and standard deviations of the mean weekly count in the new diagnosis period and control periods.	• Diagnosis period: 0-18 months • Control period: 18-24 months • Stratify by time categories: No (study period 2008-2012)	• Diagnosis period: 0-36 months • Control period: 36-48 months • Stratify by time categories: 2007-2011, 2012-2015, 2016-2019 (for gFBOT only: 2007-20018, 2009-2011, 2012-2015, 2016-2019)
	Apply the control chart rules 1-4 at each week, the lookback period stops when the rules can no longer be applied. Calculate four lookback periods for each encounter category using rules 1 to 4, rules 1 to 3, rules 1 to 2 and rule 1 only	• Rules 1 to 4, rules 1 to 3, rules 1 to 2 and rule 1 only, provide increasingly shorter lookback periods	• No differences
4. Determine each encounter category's signal strength	Determine the signal strength for each of the 4 lookback periods for each encounter and discard any encounter categories that do not reach the signal strength cut off.	• Signal strength cut off = 80% • *A priori* include abdominal ultrasound and gFOBT regardless of signal strength • Included encounters: 6 diagnosis and 13 procedure categories	• Signal strength cut off = 70% • No differences • Included encounters: 8 diagnosis and 21 procedure categories
	Define the number of lookback weeks for each eligible encounter category as the first set of rules that reach the signal strength	• Different lookback period for each encounter for the whole cohort (19 lookback periods)	• Different lookback periods for each encounter category in each time category (28x3 and 1x4 (gFOBT) lookback periods
5. Collect all eligible encounters	For each patient in the cohort, capture all eligible encounter	• Apply lookback periods to capture all encounters	• No differences

	categories and their dates that occurred within the unique lookback period for each.		
6. Add referring physician encounter date to each diagnostic procedure	For all procedure-based encounters, identify the referring physician and the date of the closest visit.	• Look back up to 18 months before diagnosis • If no referring physician or date could be identified, use the procedure date	• Look back 365 days from the encounter up to 48 months before diagnosis • No differences
7. Earliest encounter is the index contact for that patient	Define the first contact date by ordering all eligible and identified encounter categories per person from closest to diagnosis to furthest and defined the earliest encounter as the first contact date. Calculate the diagnostic interval as the number of days between the first contact date and cancer diagnosis date.	• If the first encounter was a procedure with a referring physician, then the referring physician date was defined as the first contact date. If a referring physician visit could not be identified, the procedure date was defined as the first contact date.	• No differences
8. Define the first encounter	Define the encounter category occurring on the first contact date	• Used a hierarchy to assign only one encounter to the first contact date.	• Described all encounters on the first contact date.

Supplemental Table 2a. Detailed OHIP and CIHI diagnosis codes and categories used to create the diagnostic interval

Category	OHIP codes	ICD-10 codes	OHIP DESCRIPTION
Colorectal cancer	153	C18^, C19^, C20^, C21.0^, C21.1^, C21.8	Malignant Neoplasms: Large intestine-excluding rectum
	154		Malignant Neoplasms: Rectum, rectosigmoid and anus
Other cancer	155	C17^, C22^, C26^, C78^, C79^, C80^	Malignant Neoplasms: Primary malignancy of liver (not secondary spread or metastatic disease)
	152		Malignant Neoplasms: Small intestine
	159		Malignant Neoplasms: Other and ill-defined sites within the digestive organs and peritoneum
	197		Malignant Neoplasms: Secondary neoplasm of respiratory and digestive systems
	198		Malignant Neoplasms: Metastatic or secondary malignant neoplasm, carcinomatosis
	199	C79	Malignant Neoplasms: Other malignant neoplasms
	195	C76	Malignant Neoplasms: Other ill-defined sites
	150	C15	Malignant Neoplasms: Esophagus
	151	C16	Malignant Neoplasms: Stomach
	157	C25	Malignant Neoplasms: Pancreas
	162	C34	Malignant Neoplasms: Bronchus, lung
	172	C44	Malignant Neoplasms: Melanoma of skin
	182	C55	Malignant Neoplasms: Body of uterus
	183	C57	Malignant Neoplasms: Ovary, fallopian tube, broad ligament
	189	C64, C68	Malignant Neoplasms: Kidney, other urinary organs
	188	C67	Malignant Neoplasms: Bladder
	200	C85, C83	Malignant Neoplasms: Lymphosarcoma, reticulosarcoma
	202	C96	Malignant Neoplasms: Other malignant neoplasms of lymphoid and histiocytic tissue
	203	C90	Malignant Neoplasms: Multiple myeloma, plasma cell leukemia
	204	C91	Malignant Neoplasms: Lymphoid leukemia (including lymphatic and histiocytic leukemia)
	239	D45	Neoplasms of Uncertain Behavior: Unspecified neoplasms (e.g., polycythemia vera)
	211	D00^, D01^, D13^, D37^, D38^, D48^, K317, K635, D12^, D20, D360, D369, D637, D484	Benign Neoplasms: Other parts of digestive system, peritoneum
	229		Benign Neoplasms: Other benign neoplasms
	230		Carcinoma in Situ: Digestive organs
	235		Neoplasms of Uncertain Behavior: Digestive and respiratory systems

	217	D12^, D25	Benign Neoplasms: Breast
	220	D27	Benign Neoplasms: Ovary (e.g., ovarian cyst)
	232	D04	Carcinoma in Situ: Skin
	233	D05, D06	Carcinoma in Situ: Breast and genito-urinary system
	234	D09	Carcinoma in Situ: Other
Signs and symptoms gastrointestinal tract	530	K20^, K21^, K22^, K25^, K26^, K29^, K30, K31, K52.89, K80^, K92.0, K92.1, K94,2, R06.6, R10^, R11^, R12, R13^, R14^, R15^, R18^, R19.0^, R19.1^, R19.2, R19.5, R19.7, R19.8, R63.0	Diseases of Esophagus, Stomach and Duodenum: Esophagitis, cardiospasm, ulcer of esophagus; stricture, stenosis, or obstruction of esophagus
	531		Diseases of Esophagus, Stomach and Duodenum: Gastric ulcer, with or without haemorrhage or perforation
	532		Diseases of Esophagus, Stomach and Duodenum: Duodenal ulcer, with or without haemorrhage or perforation
	535		Diseases of Esophagus, Stomach and Duodenum: Gastritis
	536		Diseases of Esophagus, Stomach and Duodenum: Hyperchlorhydria, hypochlorhydria, dyspepsia, indigestion
	537		Diseases of Esophagus, Stomach and Duodenum: Other disorders of stomach and duodenum
	574		Other Diseases of Digestive System: Cholelithiasis (gall stones) with or without cholecystitis
	787		Signs and Symptoms Not Yet Diagnosed: Anorexia, nausea and vomiting, heartburn, dysphagia, hiccough, hematemesis, jaundice, ascites, abdominal pain, melena, masses
	534	K28^	Diseases of Esophagus, Stomach and Duodenum: Stomal ulcer, gastrojejunal ulcer
	009	A09, K35^, K50^, K51^, K55.2^, K56^, K57^, K58^, K59^, K60^, K62^, K63.0, K63.1, K63.2, K63.3, L63.4, K63.8^, K63.9, K64^, K65^, K67^, K68.12, K68.19, K68.9, K91.0, K91.1, K91.85^, K91.89, K94.0^, K94.1	Intestinal Infectious Diseases: Diarrhea, gastro-enteritis, viral gastro-enteritis
	455		Diseases of Veins and Lymphatics: Haemorrhoids
	540		Other Diseases of Intestine and Peritoneum: Acute appendicitis, with or without abscess or peritonitis
	555		Other Diseases of Intestine and Peritoneum: Regional enteritis, Crohn's disease
	556		Other Diseases of Intestine and Peritoneum: Ulcerative colitis
	560		Other Diseases of Intestine and Peritoneum: Intestinal obstruction, intussusception, paralytic ileus, volvulus, impaction of intestine
	562		Other Diseases of Intestine and Peritoneum: Diverticulitis or diverticulosis of large or small intestine
	564		Other Diseases of Intestine and Peritoneum: Spastic colon, irritable colon, mucous colitis, constipation
	565		Other Diseases of Intestine and Peritoneum: Anal fissure, anal fistula
	567		Other Diseases of Intestine and Peritoneum: Peritonitis, with or without abscess
	557	K529, K37	Other Diseases of Intestine and Peritoneum: Mesenteric artery occlusion, other vascular conditions of intestine
	550	K40^	Hernia: Inguinal hernia, with or without obstruction
	552	K44^, K41, K42^, K43	Hernia: Femoral, umbilical, ventral, diaphragmatic or hiatus hernia with obstruction
	553		Hernia: Femoral, umbilical, ventral, diaphragmatic or hiatus hernia without obstruction
	566	K61	Other Diseases of Intestine and Peritoneum: Abscess of anal or rectal regions
	571	K74^, K70.3^	Other Diseases of Digestive System: Cirrhosis of the liver (e.g., alcoholic cirrhosis, biliary cirrhosis)

Signs and symptoms liver and gallbladder	573	K76^	Other Diseases of Digestive System: Other diseases of the liver
	576	K82^	Other Diseases of Digestive System: Other diseases of gallbladder and biliary ducts
Signs and symptoms, anemia	280	D50^	Iron deficiency anaemia
Signs and symptoms Haematologic conditions	281	D51^, D52^, D53^, D60^, D61^, D62, D63^, D64^	Pernicious anaemia
	282	D58	Hereditary hemolytic anaemia (e.g., thalassemia, sickle-cell anaemia)
	284		Aplastic anaemia
	285		Other anaemias
	283	D59	Acquired hemolytic anaemia, excluding hemolytic disease of newborn
	286	D68	Coagulation defects (e.g., hemophilia, other factor deficiencies)
	287	D69^	Purpura, thrombocytopenia, other hemorrhagic conditions
	289	D75^	Other diseases of blood, marrow, spleen
Signs and symptoms, nutritional and metabolic disorders	269	E56^, E58, E59, E60, E61^, E63^	Nutritional and Metabolic Disorders: Vitamin and other nutritional deficiencies
	242	E05^	Endocrine Glands: Hyperthyroidism, thyrotoxicosis, exophthalmic goitre
	259	E34	Diabetes: Other endocrine disorders
	277	E88^, R634	Nutritional and Metabolic Disorders: Other metabolic disorders
	278	E66^	Nutritional and Metabolic Disorders: Obesity
Signs and symptoms, NOS	595	G47.3^, N23, N30^, N36^, N39.0, N39.8, N39.9, R00.0, R04.0, R04.2, R05, R06.0^, K06.4, R07^, R26.0, R27.0, R31^, R33^, R35.1, R42, R51, R55, R56^, R57^, R60^, R64, R68.89, R69, R70^, R71^, R72^, R73^, R74^, R75^, R76^, R77^, R78^, R79^, N31	Diseases of the Urinary System: Cystitis
	599		Diseases of the Urinary System: Other disorders of urinary tract
	780		Non-specific Abnormal Findings: Ataxia
	785		Signs and Symptoms Not Yet Diagnosed: Chest pain, tachycardia, syncope, shock, edema, masses
	786		Signs and Symptoms Not Yet Diagnosed: Epistaxis, hemoptysis, cough, dyspnea, masses, shortness of breath, hyperventilation, sleep apnea
	788		Signs and Symptoms Not Yet Diagnosed: Renal colic, urinary retention, nocturia, masses
	790		Non-specific Abnormal Findings: Non-specific findings on examination of blood
	796		Non-specific Abnormal Findings: Other non-specific abnormal findings
	799		Non-specific Abnormal Findings: Other ill-defined conditions
	682	L03^	Cellulitis, abscess (diseases of the skin)
	685	L05^	Pilondial cyst or abscess (diseases of the skin)
	680	L02^	Boil, carbuncle, furunculosis (diseases of the skin)

	707	L89^	Decubitus ulcer, bed sore (diseases of the skin)
	781	M79.6, M79.1, M25.5, M25.4,	Signs and Symptoms Not Yet Diagnosed: Leg cramps, leg pain, muscle pain, joint pain, arthralgia, joint swelling, masses
Bacterial or viral disease	040	B96^	Other bacterial disease
	079	B33^	Other viral disease
	038	A41^, A40^	Other Bacterial Diseases: Septicemia, blood poisoning
Mental and neurological disorders	290	F03^	Psychoses: Senile dementia, presenile dementia
	300	F32^, F42^,	Neuroses and Personality Disorders: Anxiety neurosis, hysteria, neurasthenia, obsessive compulsive neurosis, reactive depression
	303	F10^	Neuroses and Personality Disorders: Alcoholism
	305	Z72.0	Neuroses and Personality Disorders: Tobacco abuse
	309	F43^	Neuroses and Personality Disorders: Adjustment reaction
	311	F33^	Neuroses and Personality Disorders: Depressive or other non-psychotic disorders, not elsewhere classified
	797	R54^, R41.8, R41.0	Non-specific Abnormal Findings: Senility, senescence
	332	G20	Parkinson's (diseases of the nervous system)
	349	G96.9, G96.8	Other diseases of the central nervous system (diseases of the nervous system)
	345	G40	Epilepsy (diseases of the nervous system)
	909	Z60^, Z63,	Social, Marital and Family Problems: Other problems of social adjustment
Circulatory system	394	I05^, I08.0, I08.3	Rheumatic Fever and Rheumatic Heart Disease: Mitral stenosis, mitral insufficiency
	402	I11^	Hypertensive Disease: Hypertensive heart disease
	410	I21^, I24^	Ischaemic and Other Forms of Heart Disease: Acute myocardial infarction
	412	I25.2, I25.82	Ischaemic and Other Forms of Heart Disease: Old myocardial infarction, chronic coronary artery disease of arteriosclerotic heart disease, without symptoms
	413	I20^	Ischaemic and Other Forms of Heart Disease: Acute coronary insufficiency, angina pectoris, acute ischaemic heart disease
	415	I26.9	Ischaemic and Other Forms of Heart Disease: Pulmonary embolism, pulmonary infarction
	426	I45	Ischaemic and Other Forms of Heart Disease: Heart blocks, other conduction disorders
	427	I48^	Ischaemic and Other Forms of Heart Disease: Paroxysmal tachycardia, atrial or ventricular flutter or fibrillation, cardiac arrest, other arrythmias
	428	I50^	Ischaemic and Other Forms of Heart Disease: Congestive heart failure
	429	I51^	Ischaemic and Other Forms of Heart Disease: All other forms of heart disease
	436	G45^, G46^, I63	Cerebrovascular Disease: Acute cerebrovascular accident, C.V.A., stroke
	437	I67^, G93.4^	Cerebrovascular Disease: Chronic arteriosclerotic cerebrovascular disease, hypertensive encephalopathy

	440	I70^	Diseases of Arteries: Generalized arteriosclerosis, atherosclerosis
	441	I71^	Diseases of Arteries: Aortic aneurysm (non-syphilitic)
	443	I73^	Diseases of Arteries: Raynaud's disease, Buerger's disease, peripheral vascular disease, intermittent claudication
	447	177^	Diseases of Arteries: Other disorders of arteries
	451	I80^	Diseases of Veins and Lymphatics: Phlebitis, thrombophlebitis
	459	I99^	Diseases of Veins and Lymphatics: Other disorders of circulatory system
Genito-urinary system	584	N17^	Diseases of the Urinary System: Acute renal failure
	585	N19, N18	Diseases of the Urinary System: Chronic renal failure, uremia
	590	N10^, N11^,	Diseases of the Urinary System: Acute or chronic pyelonephritis, pyelitis, abscess
	591	N13	Diseases of the Urinary System: Hydronephrosis
	592	N20^	Diseases of the Urinary System: Stone in kidney or ureter
	593	N28^	Diseases of the Urinary System: Other disorders of kidney or ureter
	600	N40^	Diseases of Male Genital Organs: Benign prostatic hypertrophy
	601	N41^	Diseases of Male Genital Organs: Prostatitis
	625	N39.3	Other Disorders of Female Genital Tract: Dyspareunia, dysmenorrhea, premenstrual tension, stress incontinence
	626	N92^	Other Disorders of Female Genital Tract: Disorders of menstruation
	629	N94^	Other Disorders of Female Genital Tract: Other disorders of female genital organs
Respiratory system	486	J189^, J06^	Pneumonia - all types
	491	J41^	Chronic bronchitis
	492	J43^	Emphysema
	493	J45^	Asthma, allergic bronchitis
	496	J44^, J40^	Other chronic obstructive pulmonary disease
	511	J90^, R09.1	Pleurisy with or without effusion
	518	J98.1	Atelectasis, other diseases of lung
	519	J98.11	Other diseases of respiratory system

Supplemental Table 2b. Detailed OHIP and CIHI procedure codes and categories used to create the diagnostic interval

Category	OHIP codes	CIHI CCI codes	OHIP DESCRIPTION
Cancer-related visit (consultation or chemotherapy)	G388		Diagnostic and therapeutic procedures - injections - management of special oral chemotherapy
	G345		Diagnostic and therapeutic procedures - injections or infusions - chemotherapy - complex single agent or multi-agent therapy
	G382		Diagnostic and therapeutic procedures - injections or infusions - chemotherapy - monthly telephone supervision - supervision of chemotherapy
	G381		Diagnostic and therapeutic procedures - injections or infusions - chemotherapy - standard chemotherapy - agents with minor toxicity that require physician monitoring
	A445		Medical oncology - general listings - consultation
	A443		Medical oncology - general listings - medical specific assessment
	A448		Medical oncology - general listings - partial assessment
	A441		Medical oncology - general listings - re-assessment
	A341		Radiation oncology - general listings - complex medical specific re-assessment
	A345		Radiation oncology - general listings - consultation
	A340		Radiation oncology - general listings - medical specific assessment
	A348		Radiation oncology - general listings - partial assessment
Cardiovascular-related encounter	J022	3JY20	Clinical procedures - angiography - by catheterization - selective
	J021	3IP10	Clinical procedures associated with diagnostic radiological examinations - Angiography - angiography by catheritization - abdominal, thoracic, cervical or cranial
	G690		Diagnostic and therapeutic procedures - cardiac monitoring - cardiac loop monitoring (discontinued)
	G692		Diagnostic and therapeutic procedures - cardiac monitoring - cardiac loop monitoring (discontinued)
	G297		Diagnostic and therapeutic procedures - Cardiovascular - Angiography - angiograms
	Z437		Diagnostic and therapeutic procedures - cardiovascular - cardioversion
	G298	1IJ50	Diagnostic and therapeutic procedures - Cardiovascular - Coronary angioplasty stent
	Z442		Diagnostic and therapeutic procedures - Cardiovascular - selective coronary catheritization - both arteries
	Z459	1IS53	Diagnostic and therapeutic procedures - cardiovascular - vascular cannulation - arterial punction
	G268		Diagnostic and therapeutic procedures - cardiovascular - vascular cannulation - artery
	G269		Diagnostic and therapeutic procedures - cardiovascular - vascular cannulation - artery
	Z456		Diagnostic and therapeutic procedures - cardiovascular - venipuncture - insertion of implantable central venous catheter
	Z440		Diagnostic and therapeutic procedures - Cardiovascular -haemodynamic/flow/metabolic studies - left heart - retrograde aortic
	Z434		Diagnostic and therapeutic procedures - Cardiovascular system - Transluminal coronary angioplasty - one or more
	G321	2HZ24	Diagnostic and therapeutic procedures - ECG - automatic implantable defibrillator - prof component

	G659		Diagnostic and therapeutic procedures - ECG - continuous monitoring - prof comp - 60h-13days
	G684		Diagnostic and therapeutic procedures - ECG - contiguous monitoring - tech comp - recording 60h-13days
	G685		Diagnostic and therapeutic procedures - ECG - continuous monitoring - tech comp - scanning 60h-13days
	G112		Diagnostic and therapeutic procedures - ECG - Bipyramidal Thallium stress test - prof component
	G111		Diagnostic and therapeutic procedures - ECG - Bipyramidal Thallium stress test - tech component
	G180		Diagnostic and therapeutic procedures - ECG - dual chamber reprogram - prof comp
	G181		Diagnostic and therapeutic procedures - ECG - dual chamber reprogram - tech comp
	G307		Diagnostic and therapeutic procedures - ECG - pacemaker wave analysis - prof comp
	G308		Diagnostic and therapeutic procedures - ECG - pacemaker wave analysis - tech comp
	G650		Diagnostic and therapeutic procedures - ECG - prof component - 12 to 35 hours recording
	G658		Diagnostic and therapeutic procedures - ECG - prof component - 36 to 59 hours recording
	G284		Diagnostic and therapeutic procedures - ECG - single chamber reprogram - tech comp
	G319		Diagnostic and therapeutic procedures - ECG - stress testing - max stress ECG - prof component
	G315		Diagnostic and therapeutic procedures - ECG - stress testing - max stress ECG - tech component
	G651		Diagnostic and therapeutic procedures - ECG - tech component - 12 to 35 hours recording
	G652		Diagnostic and therapeutic procedures - ECG - tech component - 12 to 35 hours screening
	G682		Diagnostic and therapeutic procedures - ECG - tech component - 36 to 59 hours recording
	G683		Diagnostic and therapeutic procedures - ECG - tech component - 36 to 59 hours scanning
	G313		Diagnostic and therapeutic procedures - ECG - twelve lead - prof comp
	G310		Diagnostic and therapeutic procedures - ECG - twelve lead - tech comp
	G577		Diagnostic and therapeutic procedures - echocardiography - cardiac doppler (discontinued)
	G570		Diagnostic and therapeutic procedures - Echocardiography - complete study - tech comp
	G575		Diagnostic and therapeutic procedures - echocardiography - focused study – prof comp
	G574		Diagnostic and therapeutic procedures - echocardiography - focused study - tech comp
	G571		Diagnostic and therapeutic procedures - Echocardiography - prof component
	G583		Diagnostic and therapeutic procedures - Echocardiography - prof component
	G582		Diagnostic and therapeutic procedures - Echocardiography - tech component
	G572		Diagnostic and therapeutic procedures - echocardiography (discontinued)
	G578		Diagnostic and therapeutic procedures - echocardiography (discontinued)
	G418		Diagnostic and therapeutic procedures - Neurology - Electroencephalography - Routine EEG - prof component
	J500		Diagnostic ultrasound - vascular laboratory fees - ankle pressure measurements - (discontinued)
	J200		Diagnostic ultrasound - vascular laboratory fees - ankle pressure measurements - 4 seg pressure recordings
	J206		Diagnostic ultrasound - vascular system - duplex evaluation of portal hypertension
	J501		Diagnostic ultrasound - vascular system - duplex scan (discontinued)
	J502		Diagnostic ultrasound - vascular system - duplex scan (discontinued)
	J201		Diagnostic ultrasound - vascular system - Extra-cranial vessel assessment - duplex scan
	J193	3KR30	Diagnostic ultrasound - vascular system - peripheral vessel assessment - doppler scan

	J202		Diagnostic ultrasound - vascular system - peripheral vessel assessment - duplex scan
	J198		Diagnostic ultrasound - vascular system - venous assessment - bilateral
	J808		Nuclear medicine - Cardiovascular system - Myocardial Perfusion Scintigraphy - delayed
Consult or assessment - Gastroenterology	A411		Gastroenterology - general listings - Complex medical specific re-assessment
	A415	None	Gastroenterology - general listings - Consultation
	A545		Gastroenterology - general listings - limited consultation
	A413		Gastroenterology - general listings - Medical specific assessment
	A414		Gastroenterology - general listings - Medical specific re-assessment
	A418		Gastroenterology - general listings - Partial assessment
	A416		Gastroenterology - general listings - repeat consultation
	C415		Gastroenterology - non-emergency hospital in-patient services - Consultation
Consult or assessment - General surgery	A035	None	General surgery - general listing - Consultation
	A034		General surgery - general listing - Partial assessment
	A036		General surgery - general listing - Repeat consultation
	A935		General surgery - general listing - special surgical consult
	A033		General surgery - general listing - Specific assessment
	C035		General surgery - non-emergency hospital in-patient services - Consultation
	C033		General surgery - non-emergency hospital in-patient services - specific assessment
	C034		General surgery - non-emergency hospital in-patient services - Specific re-assessment
Consult or assessment - Internal medicine	A145	None	Consultation if physician's practice is predominantly cardiology, respirology or gastroenterology (replaces A135)
	A131		Internal and occupational medicine - general listings - Complex medical specific re-assessment
	A130		Internal and occupational medicine - general listings - Comprehensive internal medicine consultation
	A135		Internal and occupational medicine - general listings - Consultation
	A435		Internal and occupational medicine - general listings - limited consultation
	A133		Internal and occupational medicine - general listings - Medical specific assessment
	A134		Internal and occupational medicine - general listings - Medical specific re-assessment
	A138		Internal and occupational medicine - general listings - Partial assessment
	A136		Internal and occupational medicine - general listings - repeat consultation
	C130		Internal and occupational medicine - non-emergency hospital in-patient services - comprehensive internal medicine consult
	C135		Internal and occupational medicine - non-emergency hospital in-patient services - Consultation
	C435		Internal and occupational medicine - non-emergency hospital in-patient services - limited consultation
	C133		Internal and occupational medicine - non-emergency hospital in-patient services - medical specific assessment

	C134		Internal and occupational medicine - non-emergency hospital in-patient services - medical specific re-assessment
	C136		Internal and occupational medicine - non-emergency hospital in-patient services - repeat consultation
Consult or assessment - Other specialty	C015		Anaesthesiology - Anaesthesiology consult
	A015		Anaesthesiology - Anaesthesiology consult
	A014		Anaesthesiology - Anaesthesiology partial assessment
	C013		Anaesthesiology - Anaesthesiology specific assessment
	A013		Anaesthesiology - Anaesthesiology specific assessment
	C215		Anaesthesiology - Limited consultation acute pain management
	A215		Anaesthesiology - Limited consultation pain management
	A601		Consultations and visits - Cardiology - general listings - complex medical specific re-assessment
	A605		Consultations and visits - Cardiology - general listings - consultation
	C605		Consultations and visits - Cardiology - general listings - consultation
	A603		Consultations and visits - cardiology - general listings - medical specific assessment
	A604		Consultations and visits - Cardiology - general listings - medical specific re-assessment
	A608		Consultations and visits - Cardiology - general listings - partial assessment
	A606		Consultations and visits - Cardiology - general listings - repeat consultation
	A155		Consultations and visits - Endocrinology and metabolism - general listings - consultation
	A153		Consultations and visits - Endocrinology and metabolism - general listings - medical specific assessment
	A151		Consultations and visits - Endocrinology and metabolism - general listings - medical specific re-assessment
	A154		Consultations and visits - Endocrinology and metabolism - general listings - medical specific re-assessment
	A905		Consultations and visits - family practice - general listings - special palliative care assessment - limited consultation
	K023		Consultations and visits - family practice - palliative care support
	K032		Consultations and visits - Family practice - specific neurocognitive assessment
	A645		Consultations and visits - General thoracic surgery - general listings - consultation
	A644		Consultations and visits - General thoracic surgery - general listings - partial assessment
	A075		Consultations and visits - Geriatrics - general listings - consultation
	A770		Consultations and visits - geriatrics - general listings - extended comprehensive geriatric consultation
	A073		Consultations and visits - Geriatrics - general listings - medical specific assessment
	C075		Consultations and visits - geriatrics - non-emergency hospital in-patient services - consultation
	A611		Consultations and visits - Haematology - general listings - complex medical specific re- assessment
	A615		Consultations and visits - Haematology - general listings - consultation
	A613		Consultations and visits - Haematology - general listings - medical specific assessment
	A614		Consultations and visits - Haematology - general listings - medical specific re-assessment

	A618		Consultations and visits - Haematology - general listings - partial assessment
	C615		Consultations and visits - Haematology - non-emergency hospital in-patient services - consultation
	A161		Consultations and visits - Nephrology - general listings - complex medical specific re-assessment
	A165		Consultations and visits - Nephrology - general listings - consultation
	A163		Consultations and visits - Nephrology - general listings - medical specific assessment
	A168		Consultations and visits - Nephrology - general listings - partial assessment
	A181		Consultations and visits - neurology - general listings - complex medical specific assessment
	A185		Consultations and visits - neurology - general listings - consultation
	A183		Consultations and visits - neurology - general listings - medical specific assessment
	A205		Consultations and visits - obstetrics and gynaecology - consultation
	A203		Consultations and visits - obstetrics and gynaecology - specific assessment
	A481		Consultations and visits - rheumatology - general listings - complex medical specific re-assessment
	A483		Consultations and visits - rheumatology - general listings - medical specific assessment
	A355		Consultations and visits - Urology - general listings - consultation
	A354		Consultations and visits - Urology - general listings - partial assessment
	A353		Consultations and visits - Urology - general listings - specific assessment
	C355		Consultations and visits - Urology - non-emergency hospital in-patient services - consultation
	C945		Family practice - In hospital palliative care consult
	A945		Family practice - specialist palliative care consult
	G512		Palliative care - Palliative care case management
Critical care	G557		Diagnostic and therapeutic procedures - critical care - comprehensive care - first day
	G400		Diagnostic and therapeutic procedures - critical care - ICU - physician in charge - first day
	G522		Diagnostic and therapeutic procedures - critical care - life threatening - 1/2 hour
	G521		Diagnostic and therapeutic procedures - critical care - life threatening - first 1/4 hour
	G523		Diagnostic and therapeutic procedures - critical care - life threatening - second 1/4 hour
	G395		Diagnostic and therapeutic procedures - critical care - other critical care - first 1/4 hour
	G405	1GZ31	Diagnostic and therapeutic procedures - critical care - Ventilatory support ICU - physician in charge - first day
CT - abdomen, chest, pelvis	X410	3NM20	Diagnostic radiology - CT - abdomen - With IV contrast
	X126	3OT20	Diagnostic radiology - CT - abdomen - with or without IV contrast
	X409		Diagnostic radiology - CT - abdomen - without IV contrast
	X234		Diagnostic radiology - CT - CT colonography
	X233		Diagnostic radiology - CT - Pelvis - with and without contrast
	X232		Diagnostic radiology - CT - Pelvis - with IV contrast
	X231		Diagnostic radiology - CT - Pelvis - without IV contrast

	X125	3GY20	Diagnostic radiology - CT - Thorax - with and without IV contrast
	X407		Diagnostic radiology - CT - Thorax - with IV contrast
	X406		Diagnostic radiology - CT - Thorax - without IV contrast
		3ZZ20	
CT - head, neck and spine	X408	3ER20	Diagnostic radiology - CT - complex head - with and without IV contrast
	X405	3AN20	Diagnostic radiology - CT - complex head - with IV contrast
	X402		Diagnostic radiology - CT - Complex head - without IV contrast
	X188		Diagnostic radiology - CT - Head - with and without contrast IV
	X401		Diagnostic radiology - CT - Head - with IV contrast
	X400		Diagnostic radiology - CT - Head - without IV contrast
	X404		Diagnostic radiology - CT - neck - with IV contrast
	X168	3SC20	Diagnostic radiology - CT - Spine - CT guidance of biopsy
	X417		Diagnostic radiology - CT - Spine - three-dimensional CT acquisition sequencing
	X415		Diagnostic radiology - CT - Spine - without IV contrast
Endoscopy - lower GI	547 (dx code)	2NM70^^	Diseases of Esophagus, Stomach and Duodenum: Colon Family history of colon cancer
	545 (dx code)		Diseases of Esophagus, Stomach and Duodenum: Colon Positive Fecal Occult Blood
	548 (dx code)		Diseases of Esophagus, Stomach and Duodenum: Colon Screening
	546 (dx code)		Diseases of Esophagus, Stomach and Duodenum: Colon Surveillance
	Z555		Intestines - Colonoscopy for diagnosis or ongoing management - Absence of signs or symptoms or risk factors, 50 years of age or older - sigmoid to descending colon
	Z498		Intestines - Colonoscopy for diagnosis or ongoing management - Follow up of abnormal colonoscopy - sigmoid to descending colon
	Z495		Intestines - Colonoscopy for diagnosis or ongoing management - follow up of unsatisfactory colonoscopy
	Z494		Intestines - Colonoscopy for diagnosis or ongoing management - hereditary or other bowel disorders associated with increased risk of malignancy
	Z496		Intestines - Colonoscopy for diagnosis or ongoing management - Presence of signs or symptoms - sigmoid to descending colon
	Z499		Intestines - Colonoscopy for risk evaluation - Absence of signs or symptoms, family history associated with an increased risk of malignancy – sigmoid to descending colon
	Z497		Intestines - Colonoscopy for risk evaluation - confirmatory colonoscopy - sigmoid to descending colon
	Z492		Intestines - Colonoscopy for risk evaluation - five year follow up of normal colonoscopy, absence of intervening signs or symptoms - sigmoid to descending colon
	Z580		Intestines (except rectum) - Sigmoidoscopy (using 60 cm. flexible endoscope)
	Z571	1NM87BA	Polypectomy - Excision of first polyp greater than or equal to 3mm through colonoscope

	Z765	1NQ87BA^	Polypectomy - Excision of obstructive tumour or stricture through colonoscopy 2 cm or greater
	Z491	1NM87DA	Polypectomy - FOLLOW UP OF INCOMPLETE POLYP RESECTION
	Z570	1NQ87DA	Polypectomy - Fulguration of first polyp through colonoscope
	Z543	2NQ70^^	Rectum - Endoscopy Anoscope (proctoscopy)
	Z536	2NM70BABG	Rectum - Endoscopy Sigmoidoscopy with or without anoscopy with biopsy(ies)
	Z535	2NM70BABH	Rectum - Endoscopy Sigmoidoscopy with or without anoscopy with rigid scope
Endoscopy - Non-GI	Z606	2PM70	Bladder - Endoscopy - Diagnostic with or without urethroscopy
	Z321	2GE70	Larynx - Laryngoscopy (discontinued)
	Z327	2GM70	Trachea and Bronchi - Bronchoscopy - flexible or rigid, with or without biopsy, suction or injection of contrast
Endoscopy - Upper GI	Z560	2NK70^^	Intestines - endoscopy - duodenoscopy
	Z584		Intestines - endoscopy - small bowel push entersocopy
	Z399		Oesophagus - endoscopy - gastroscopy, with or without duodenoscopy - elective
	Z400		Oesophagus - endoscopy - gastroscopy, with or without duodenoscopy -for active bleeding
	Z515	2NA70^^	Oesophagus - endoscopy - Esophagoscopy with or without biopsies
	Z527	2NF70^^	Stomach - endoscopy - gastroscopy - may include biopsies, photography and removal of polyps less than or equal to 1
Family physician - Emergency visit	H055		Consultations and visits - Emergency medicine - Emergency department physician on duty - consultation
	A888		Consultations and visits - family practice - emergency department equivalent, partial assessment
	H102		Consultations and visits - family practice - emergency department physician - daytime - comprehensive assessment
	H103		Consultations and visits - family practice - emergency department physician - daytime - multiple systems assessment
	H101		Consultations and visits - family practice - emergency department physician - daytime - re-assessment
	H104		Consultations and visits - family practice - emergency department physician - daytime - re-assessment
	H132		Consultations and visits - family practice - emergency department physician - evenings - comprehensive assessment
	H131		Consultations and visits - family practice - emergency department physician - evenings - minor assessment
	H133		Consultations and visits - family practice - emergency department physician - evenings - multiple systems assessment
	H134		Consultations and visits - family practice - emergency department physician - evenings - re-assessment
	H122		Consultations and visits - family practice - emergency department physician - nights - comprehensive assessment
	H121		Consultations and visits - family practice - emergency department physician - nights - minor assessment
	H123		Consultations and visits - family practice - emergency department physician - nights - multiple systems assessment
	H124		Consultations and visits - family practice - emergency department physician - nights - re-assessment
	H152		Consultations and visits - family practice - emergency department physician - weekend - comprehensive assessment
	H151		Consultations and visits - family practice - emergency department physician - weekend - minor assessment
	H153		Consultations and visits - family practice - emergency department physician - weekend - multiple systems assessment

	H154		Consultations and visits - family practice - emergency department physician - weekend - re-assessment
	H065		Consultations and visits - Family practice - Emergency department physician - consultation in emergency medicine
	H105		Consultations and visits - Family practice - Emergency department physician - in-patient interim admission orders
	C933		Consultations and visits - Family practice - non-emergency hospital in-patient services - on-call admission assessment (has to be non-elective)
	A933		Consultations and visits - Family practice - on-call admission assessment (has to be non-elective)
Family physician - non-emergency visit	A900		Consultations and visits - Family practice - complex house call assessment
	K033		Consultations and visits - family practice - counselling - individual support
	K030		Consultations and visits - Family practice - diabetic management fee (all-inclusive service)
	A005		Consultations and visits - Family practice - general listing - Family practice - consultation
	A004		Consultations and visits - Family practice - general listing - Family practice - general re-assessment
	A007		Consultations and visits - Family practice - general listing - Family practice - intermediate assessment or well-baby care
	A003		Consultations and visits - Family practice - general listing - general assessment
	A001		Consultations and visits - Family practice - general listing - minor assessment
	K071		Consultations and visits - family practice - home care supervision - acute home care supervision
	K072		Consultations and visits - family practice - home care supervision - chronic home care supervision
	A901		Consultations and visits - Family practice - house call assessment (discontinued)
	C003		Consultations and visits - Family practice - non-emergency hospital in-patient services - general assessment
	C004		Consultations and visits - family practice - non-emergency hospital in-patient services - general re-assessment
	C005		Consultations and visits - Family practice - non-emergency in-patient services - consultation
	C905		Consultations and visits - family practice - non-emergency in-patient services - limited consultation
	W102		Consultations and visits - Family practice - non-emergency long-term care in-patient services - Type 1 admission assessment
	K131		Consultations and visits - Family practice - periodic health visit adult aged 18-64
	K132		Consultations and visits - Family practice - periodic health visit adult aged 65+
	K015		Family practice - counselling - Counselling of relatives on behalf of terminally ill patient
	K040		Family practice - counselling - Group counselling - 2 or more persons
	K013		Family practice - Counselling - individual care
	K002		Family practice - interviews - Interview with relative authorized to make treatment decision
	K005		Family practice - primary mental health care - individual care
	K004		Family practice - Psychotherapy - 2 or more family members at the same time
	K007		Family practice - Psychotherapy - individual care
	917		Other: Annual health examination adolescent/adult Well Vision Care (with feecode A001, A003, A004, A005, A007, A901)
gFOBT, FIT	L181		Occult blood
	L179		Occult blood

	Q152	None	FOBT Completion Fee (Once per patient every two years)
	Q150		FOBT Distribution and Counselling Fee (Once per patient every two years)
	G004		Occult blood
	Q005		Preventative care, CRC screening (discontinued)
Lung-related encounter	J304		Pulmonary function studies - flow volume loop - volume versus flow study
	J306	2GT21	Pulmonary function studies - functional residual capacity - airways resistance by plethysmography
	J307		Pulmonary function studies - functional residual capacity - by body plethysmography
	J311		Pulmonary function studies - functional residual capacity - by gas dilution method
	J310		Pulmonary function studies - functional residual capacity - carbon monoxide diffusing capacity
	J323	2GZ58	Pulmonary function studies - oxygen saturation - by oximetry at rest
	J332		Pulmonary function studies - oxygen saturation - by oximetry at rest and exercise
	J319		Pulmonary function studies - Stage II - blood gas analysis
Miscellaneous	G860		Diagnostic and therapeutic procedures - dialysis - chronic dialysis - hospital haemodialysis (n=200)
	X146		Diagnostic radiology - bone mineral density - baseline test (n=300)
	Z770	2RM71	Female genital procedures - corpus uteri - incision - endometrial sampling
	S745		Female genital procedures - vagina - combined abd-vag procedure for stress incontinence - two surgeons - abdominal surgeon
	Z408	2WY71	Haematic and lymphatic surgical procedures - spleen and marrow - bone marrow core biopsy
	Z119		Integumentary system surgical procedures - skin - cryotherapy treatment of pre-malignant keratosis
	Z712	2QT71	Male genital surgical procedures - prostate - incision - biopsy
MRI - abdomen, pelvis	X451	3OT40	MRI - Abdomen - multislice sequence
	X455		MRI - Abdomen - repeat (another plane, different pulse sequence - to a maximum of 3 repeats
	X461		MRI - pelvis - multislice sequence
	X465		MRI - pelvis - repeat (another plane, different pulse sequence - to a maximum of 3 repeats
MRI - head, spine	X421	3ER40	MRI - head - multislice sequence
	X425		MRI - head - repeat
MRI - head, spine	X499	3SC40	MRI - complex spine - three dimensional
	X487		MRI - complex spine - when gadolinium is used
	X490		MRI - limited spine - multislice sequence
	X492		MRI - limited spine - repeat
Other abdominal procedure	S340		Digestive system surgical procedures - abdomen peritoneum and omentum - congenital diaphragmatic hernia - ventral - post operative
	Z594		Digestive system surgical procedures - abdomen peritoneum and omentum - incision - paracentesis abscess
	Z591	1OT52	Digestive system surgical procedures - abdomen peritoneum and omentum - paracentesis - aspiration
	S205	1NV89	Digestive system surgical procedures - Appendectomy

	S287	1OD89	Digestive system surgical procedures - excision - Cholecystectomy
	S323		Digestive system surgical procedures - Herniotomy - adolescents and adults unilateral with exploration of other side
	Z551	2OA71	Digestive system surgical procedures - liver - incision - biopsy
	Z546		Digestive system surgical procedures - rectum - excision - baron ligations (for hemorrhoids)
Surgical resection of colon	S312	1NM89^^	Abdomen, peritoneum, and omentum - incision - Laparotomy, with or without biopsy or for Hirschsprung's disease (except biopsies of stomach, liver, pancreas and multiple para-aortic lymph nodes
	S314	1NM91^^	Abdomen, peritoneum, and omentum - incision - Peritoneal abscess - abdominal
	S157	1NQ89^^	Intestines - Enterotomy - Colostomy
	S149	1NM87LA	Intestines - Enterotomy - Ileostomy
	S180	1NM87DF	Intestines - Intestinal obstruction (mechanical) - with enterotomy
	S177	1NM87RN	Intestines - Intestinal obstruction (mechanical) - with resection
	S175	1NM87DE	Intestines - Intestinal obstruction (mechanical) - without resection
	S167	1NM87RD	Intestines - Resection with anastomosis - large intestine -any portion
	S171	1NM87DN	Intestines - Resection with anastomosis - Left hemicolectomy with anterior resection or proctosigmoidectomy (anastomosis below peritoneal reflection & mobilization of splenic flexure
	S165	1NM87RE	Intestines - Resection with anastomosis - other
	S166	1NM87DX	Intestines - Resection with anastomosis - Small and large intestine terminal ileum, cecum and ascending colon (right hemicolectomy)
	S169	1NM87TF	Intestines - Resection with anastomosis - Total colectomy with ileo-rectal anastomosis
	S184	1NM87DY	Intestines - suture - Suture of intestine
	S217	1NM87TG	Rectum - proctectomy - 2 surgeon team - Hartmann procedure
	S213	1NQ87LA	Rectum - proctectomy - Anterior resection or proctosigmoidectomy (anastomosis below peritoneal reflection)
		1NM59^^	Destruction, large intestine
		1NM59^^	Destruction, large intestine
		1NQ87CA	Excision partial, rectum, per orifice approach, closure by apposition technique or no closure required
		1NQ87PF	Excision partial, rectum, posterior approach, closure by apposition technique or no closure required
		1NQ87RD	Excision partial, rectum, open abdominal, colorectal anastomosis
		1NQ87DE	Excision partial, rectum, endoscopic, colorectal anastomosis
		1NQ87PB	Excision partial, rectum, per orifice approach, colorectal anastomosis
		1NQ87TF	Excision partial, rectum, open abdominal, colostomy with closure of rectal stump or submucosa fistula
		1NQ87DX	Excision partial, rectum, endoscopic, approach, colostomy with closure of rectal stump or submucosa fistula
Ultrasound - Abdomen, pelvis	J435	3OT30^^	Diagnostic ultrasound - abdomen and retroperitoneum - complete study
	J428		Diagnostic ultrasound - abdomen and retroperitoneum - limited study
	J162		Diagnostic ultrasound - Pelvis - complete

	J462		Diagnostic ultrasound - pelvis - complete study
	J138		Diagnostic ultrasound - Pelvis - Intracavitary ultrasound (e.g., transrectal, transvaginal
	J463		Diagnostic ultrasound - pelvis - limited study - for other than pregnancy
	J163		Diagnostic ultrasound - Pelvis - limited study -for other than pregnancy
	J438		Diagnostic ultrasound - pelvis - transrectal/vaginal
	J135		Diagnostic ultrasound - thorax, abdomen and retroperitoneum - Abdominal scan - complete
	J128		Diagnostic ultrasound - thorax, abdomen and retroperitoneum - Abdominal scan - limited study (e.g., gallbladder only, aorta only or follow-up study)
Ultrasound - other	J105		Diagnostic ultrasound - head and neck - face and/or neck
	J405		Diagnostic ultrasound - head and neck - face and/or neck (discontinued)
	J127	3YM30	Diagnostic ultrasound - miscellaneous - breast- scan B-mode
	J182		Diagnostic ultrasound - miscellaneous - extremities
	J183	3QC30	Diagnostic ultrasound - miscellaneous - scrotal
	J125	3GY30	Diagnostic ultrasound - thorax abdomen and retroperitoneum - thorax - chest masses, pleural effusion
Xray - abdomen, chest, pelvis	X100	3NL10	Diagnostic radiology - chest and abdomen - Abdomen - single view
	X101	3OT10	Diagnostic radiology - chest and abdomen - Abdomen - two or more views
	X090	3GY10	Diagnostic radiology - chest and abdomen - chest - single view
	X092		Diagnostic radiology - chest and abdomen - chest - three or more views
	X091		Diagnostic radiology - chest and abdomen - chest - two view
	X039	3SL10	Diagnostic radiology - chest and abdomen - Ribs - two or more view
	X197	3OT12	Diagnostic radiology - Fluoroscopy - Abdomen
	X195		Diagnostic radiology - fluoroscopy - chest
	X113	3NQ10	Diagnostic radiology - gastrointestinal tract - colon - air contrast, primary or secondary, including survey films, if taken
	X112	3NZ10	Diagnostic radiology - gastrointestinal tract - colon - barium enema including survey film, if taken
	X104	3NM10	Diagnostic radiology - gastrointestinal tract - Oesophagus, stomach and duodenum - double contrast, including survey film, if taken
	X103	3NA10	Diagnostic radiology - gastrointestinal tract - Oesophagus, stomach and duodenum - double contrast, including survey film, if taken, and small bowel
	X111	3NK10	Diagnostic radiology - gastrointestinal tract - small bowel only - when only examination is performed during patients visit
	X036	3SQ10	Diagnostic radiology - spine and pelvis - Pelvis and/or hip(s) - one view
	X038	3VA10	Diagnostic radiology - spine and pelvis - Pelvis and/or hip(s) - three+ view
	X037		Diagnostic radiology - spine and pelvis - Pelvis and/or hip(s) - two view
Xray - other	X016	3DL10	Diagnostic radiology - head and neck - mastoid - eye for foreign body
	X106		Diagnostic radiology - gastrointestinal tract - pharynx and esophagus - cine or videotape

	X227	3WA10	Diagnostic radiology - lower extremities - ankle - four views
	X063	3VZ10	Diagnostic radiology - lower extreme tires - femur including one joint - two views
	X060	3VG10	Diagnostic radiology - lower extremities - hip - two views
	X066	3VQ10	Diagnostic radiology - lower extreme tires - tibia and fibula including one joint - two views
	X178		Diagnostic radiology - miscellaneous exams - mammogram no signs or symptoms
	X194		Diagnostic radiology - miscellaneous exams - mammogram no signs or symptoms
	X050	3TK10	Diagnostic radiology - upper extremities - humerus including one joint - two views
	X212	3TA10	Diagnostic radiology - upper extremities - shoulder - three views
	X048		Diagnostic radiology - upper extremities - shoulder - two views
Xray - spine	X025	3SC10	Diagnostic radiology - spine and pelvis - cervical spine - two or three views
	X205		Diagnostic radiology - spine and pelvis - Lumbar or lumbosacral spine - four or five views
	X206		Diagnostic radiology - spine and pelvis - Lumbar or lumbosacral spine - six or more views
	X028		Diagnostic radiology - spine and pelvis - Lumbar or lumbosacral spine - two or three views
	X035		Diagnostic radiology - spine and pelvis - Sacro-iliac joints - two or three views
	X204		Diagnostic radiology - spine and pelvis - thoracic spine - three or more views
	X027		Diagnostic radiology - spine and pelvis - thoracic spine - two views
Biopsy		2NM71^^	Biopsy, large intestine
		2NF71	Biopsy, stomach
		2NK71^^	Biopsy, small intestine
		2NK71^^	Biopsy, small intestine
		2NQ71	Biopsy, rectum
		2NA71	Biopsy, esophagus
	Z340	2GT71	Respiratory surgical procedures - lungs and pleura - incision - biopsy of lung

Chapter 6 : Manuscript 3. Income inequalities in time to colon cancer diagnosis

6.1 Preface

Shortening the diagnostic interval for patients experiencing low income could improve outcomes, such as reduced wait times for treatment, improved survival, and possibly decreasing inequalities later in the cancer continuum. In this manuscript, I used the definition for the diagnostic interval defined in Chapter 5 and examined inequalities in the colon cancer diagnostic interval and pathways by neighbourhood income.

While the first half of this thesis identifies that neighbourhood income is not a good proxy for individual income, unfortunately, the Statistics Canada data is not linked to physician billing data needed to define the diagnostic interval, and the ICES data that contains physician billing data does not have individual-level income measures. Moreover, I could not perform a probabilistic bias analysis to obtain bias-adjusted effects of income on the diagnostic interval since I could not obtain the bias parameters by the outcome. Given the heterogeneity within neighbourhoods in Ontario and the difficulty of interpreting neighbourhood income quintiles as place-based measures, I interpret neighbourhood income as a proxy for individual income in this chapter. I accept the limitations of this and stress that results are likely attenuated. The findings from this chapter contribute to the knowledge of inequalities in the diagnostic phase of colon cancer and highlight the need for access to individual income measures.

This manuscript has been submitted to *Cancer Medicine.*

Citation: Davis LE, Strumpf EC, Patel SV, Mahar AL. Income inequalities in time to colon cancer diagnosis. Submitted to *Cancer Medicine*.

6.2 Manuscript 3

ABSTRACT

Introduction. Patients experiencing low income have worse outcomes throughout the cancer care continuum, however, little is known about income and the diagnostic interval. We described diagnostic pathways by neighbourhood income and investigated the association between income and the diagnostic interval.

Methods. This was a retrospective cohort study of colon cancer patients diagnosed 2007-2019 in Ontario using routinely collected data. The diagnostic interval was defined as the number of days from the first colon cancer encounter to diagnosis. Asymptomatic pathways were defined as first encounter with a colonoscopy or guaiac fecal occult blood test not occurring in the emergency department and were examined separately from symptomatic pathways. Quantile regression was used to determine the association between neighbourhood income quintile and the conditional 50^{th} and 90^{th} percentile diagnostic interval controlling for age, sex, rural residence, and year of diagnosis.

Results. 64,303 colon cancer patients were included. Patients residing in the lowest income neighbourhoods were more likely to be diagnosed through symptomatic pathways and in the emergency department. Living in low-income neighbourhoods was associated with longer 50^{th} and 90th-percentile symptomatic diagnostic intervals compared to patients living in the highest income neighbourhoods. For example, the 90^{th} percentile diagnostic interval was 15 days (95% CI 6-23) longer in patients living in the lowest income neighbourhoods compared to the highest.

Conclusion. These findings reveal income inequities during the diagnostic phase of colon cancer. Future work should determine pathways to reducing inequalities along the diagnostic interval and evaluate screening and diagnostic assessment programs from an equity perspective.

INTRODUCTION

Early detection of colon cancer through organized screening programs and efficient diagnostic pathways is critical for improving overall survival.[1,2] Studies have shown that prolonged diagnostic intervals can result in adverse outcomes, including later stage at diagnosis, increased patient anxiety, and worse survival rates.[3–5] Defined as the time from screening or initial presentation of symptoms to the cancer diagnosis, the diagnostic interval is a modifiable factor that can improve adverse outcomes.[6,7]

Barriers to navigating the healthcare system for a cancer diagnosis exist at the patient, provider and health system levels.[8] These barriers are prevalent for all patients, but they disproportionately effect individuals facing structural inequities, such as poverty, leading to inequalities in cancer outcomes.[9] Cancer patients experiencing low income are more likely to have worse stage at diagnosis, and lower rates of survival and screening, however, little is known about income inequalities in the diagnostic interval specifically.[10–16] To our knowledge, only one other study has examined the colon cancer diagnostic interval by income, reporting a median diagnostic interval of 6.5 days longer in patients living in low-income neighbourhoods compared to high-income neighbourhoods, however, the objective of this study was to examine multiple factors associated with the diagnostic interval and not income specifically.[16]

Thus, the objective of this work was to describe diagnostic interval characteristics by neighbourhood income quintile and estimate the associations between neighbourhood income and the length of diagnostic interval. By conducting this research, we seek to contribute to the understanding of income inequalities during the diagnostic interval and subsequently cancer outcomes. The findings of this study will provide valuable insights for developing and evaluating targeted interventions to reduce inequities in diagnostic care and improve patient outcomes.

METHODS

Study design

This was a population-based retrospective cohort study using linked routinely collected administrative healthcare databases held at ICES (formerly the Institute for Clinical Evaluative Sciences) in Ontario, Canada. Ethics approval was obtained from McGill University Research Ethics Board (#A04-M37-22A), and we followed privacy guidelines set out by ICES.

Data sources

Data were obtained from data holdings at ICES which houses data on all publicly funded healthcare interactions in Ontario, including cancer diagnostic procedures and investigations. Databases were linked using unique encoded identifiers and analyzed through the ICES remote desktop. Datasets are described in detail in Supplemental Table 1. Briefly, we used the Ontario Cancer Registry (OCR)[17], hospitalization data from the Canadian Institute of Health Information (CIHI) Same Day Surgery (SDS) and Discharge Abstract Database (DAD)[18], emergency department data from the National Ambulatory Care Reporting System (NACRS)[19], physician claims data from Ontario Health Insurance Plan (OHIP) billing dataset and demographic data from the Registered Persons Database (RPDB).

Study population

All residents of Ontario have universal, publicly funded health insurance, including primary and cancer care coverage, through a government-administered single-payer system. The study included Ontarian adults with a first colon cancer diagnosis (International Classification of

Diseases for Oncology (ICD-O-3) codes C18.0, C18.2-C18.9) registered between January 1, 2007, and December 31, 2019, in the Ontario Cancer Registry. We excluded individuals who had a death date before their diagnosis date, those diagnosed with multiple cancers on the same day, no Ontario Health Insurance Plan (OHIP) eligibility two years before the diagnosis, those for whom the first contact encounter for the diagnostic interval could not be identified and who had missing information on income.

Measures

Exposure. While we sought out to understand the effect of individual income on the diagnostic interval, measuring the diagnostic interval must occur in provincial ICES datasets where physician billing information is linked to hospital records and the cancer registry. Currently, these data are not linked to individual measures of income. As a result, we used neighbourhood income as a proxy for individual income while understanding the limitations of this approach.[20] In the absence of individual data, neighbourhood measures are commonly used to approximate individual income in cancer studies.[20,21] Neighbourhood income quintiles were obtained from the RPDB and measured using the Postal Code Conversion File (PCCF+) linked to the postal code at diagnosis. The PCCF+ neighbourhood income variable is created by Statistics Canada using census summary data and represents the median, before-tax, household-adjusted income within each dissemination area.[22] Dissemination areas are Statistics Canada's smallest geographical unit representing approximately 400-700 individuals per area.[23] Quintiles are created by ranking dissemination areas within each census metropolitan area (CMA), census agglomeration (CA) or other region from lowest to highest, then dividing into fifths. Individuals in quintile 1 reside in neighbourhoods with the lowest income and individuals in quintile 5 reside in neighbourhoods with the highest income.

Outcome. Following the Aarhus statement, we defined the diagnostic interval as the number of days from the earliest healthcare encounter (physician visit or hospital admission) related to colon cancer to the diagnosis date, usually the first malignant biopsy date.[6] We modified established methods from Groome et al. and Webber et al.[24,25] to define the earliest healthcare encounter using different lookback periods for each encounter category. These methods are described in detail elsewhere and have been used in CRC and breast cancer.[24,26,27] Briefly, we identified and categorized encounters occurring more frequently in the 0-3 months compared to the 24-27 months before diagnosis and determined cancer-related lookback periods for each encounter category using statistical process control.[28] We identified referring physician visits for all procedure-based encounters as the first visit with that referring physician that occurred less than 365 days from the procedure date. The earliest encounter was defined as the first eligible healthcare encounter, and we calculated the diagnostic interval as the number of days between the first encounter date or referring physician date to the diagnosis date. Our modification included extending the lookback period to 2 years to identify encounters, including all encounters that demonstrate an increase in the 0-3 months before diagnosis regardless of relation to colon cancer, and using a more liberal cut-off for the statistical process control.

Diagnostic interval characteristics. Other diagnostic interval variables are described in detail in Supplemental Table 2. Variables describing the diagnostic interval were measured along the diagnostic interval and included: first encounter type, symptomatic or asymptomatic pathway, referring physician as first contact, receipt of lower gastrointestinal (GI) endoscopies, number of visit days, and the summary of the diagnostic pathway. The first encounter type was defined as the earliest category of encounter that occurred on the first encounter date; patients could have more than one encounter on their first encounter date. An asymptomatic pathway was

defined as an interval where the first encounter was a guaiac fecal occult blood test (gFOBT) or lower GI endoscopy that occurred alone or in combination with a consultation and did not occur in the ED. An interval was considered symptomatic if there was a symptom-related encounter or non-screening procedure as the first encounter or if the first encounter occurred in the ED. The diagnostic pathway was summarized in 9 possible pathways: 1) asymptomatic; 2) lower GI endoscopy alone, presenting in the emergency department (ED); 3) lower GI endoscopy alone, not presenting in the ED; 4) lower GI endoscopy and imaging presenting in the ED; 5) lower GI endoscopy and imaging not presenting in the ED; 6) imaging alone, presenting in the ED; 7) imaging alone, not presenting in the ED; 8) no lower GI endoscopy or imaging presenting in the ED; 9) no lower GI endoscopy or imaging not presenting in the ED.[29]

Patient characteristics. Covariates are detailed in Supplemental Table 2. Demographic and cancer-related variables were measured in the year of diagnosis. Comorbidities were measured using the Elixhauser comorbidity index, which measured hospitalizations two years before cancer diagnosis and was dichotomized as ≥4 and <4.[30] Rural residence was measured by linking postal codes at the time of diagnosis to the Rurality Index of Ontario (RIO), a function of population size, distance to family practitioners and travel time to access healthcare.[31] RIO values were dichotomized as ≥45 for rural and <45 for urban residences.[31] Histology and TNM stage at diagnosis were obtained from the Ontario Cancer Registry. Stage at diagnosis represents the best International Union for Cancer Control and American Joint Committee on Cancer (UICC/AJCC) stage, a combination of the Collaborative Staging approach and data from medical records at regional cancer centres. Stage was broadly categorized as stages I/II/III/IV and unknown. Histology was dichotomized as adenocarcinoma and non-adenocarcinoma using morphology codes.

Statistical Analysis

We described the cohort demographics and disease characteristics by neighbourhood income quintile. The diagnostic interval in days and its characteristics were described by symptom status and neighbourhood income quintile. Means, medians and interquartile ranges were presented for continuous variables and numbers and proportions for categorical variables. Chi-square tests were used to test significant differences between categorical variables and neighbourhood income quintiles.

Quantile regression was used to estimate the association between neighbourhood income quintile and the conditional median diagnostic interval, stratified by symptomatic and asymptomatic pathways.[32] Quantile regression is useful in situations where the outcome is left-skewed, such as with the diagnostic interval and allows us to examine inequalities at each percentile.[32] We present effect estimates at the 50^{th} and 90^{th} percentile to understand the association of neighbourhood income at the median diagnostic interval and the 90^{th} percentile diagnostic interval where patients have the longest intervals and are most at risk for poor outcomes. Estimates at the 10^{th} percentile were initially explored to understand the association of income for the patients with the shortest diagnostic intervals but were not presented due to a lack of variation. We present 95% confidence intervals (CI) and p-values <0.05 represent statistical significance. Multivariable models included continuous age, sex, rural residence, and year of diagnosis. Comorbidities and stage at diagnosis were conceptualized as being on the causal pathway and, therefore, not included in the multivariable models.[33] We performed an additional analysis stratifying models by stage at diagnosis to determine any differences in the association between neighbourhood income and the diagnostic interval at different stages. SAS version 9.4 was used for all analyses.

RESULTS

Study cohort

67,428 individuals were diagnosed with colon cancer between 2007 and 2019. 3,126 were excluded for a final cohort of 64,303 patients (Figure 1). There were some demographic and cancer differences by neighbourhood income quintile (Table 1). The median age at diagnosis ranged from 73 (IQR 63-81) for individuals living in neighbourhoods with the lowest income and 71 (IQR 62-81) for individuals in the highest neighbourhood income quintile. Individuals residing in the lowest income neighbourhoods were more likely to be female, have more comorbidities, live in rural areas, have missing stage and less likely to be diagnosed at stage 1 compared to individuals living in the highest income neighbourhoods.

FIGURE 6.1 COHORT EXCLUSIONS

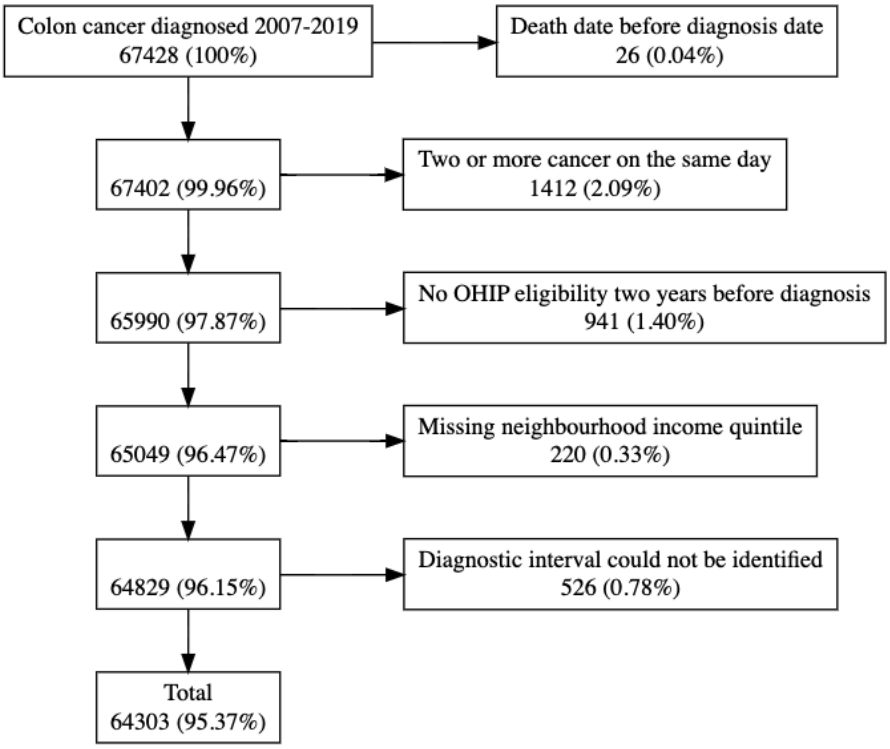


TABLE 6.1. DEMOGRAPHIC AND DISEASE CHARACTERISTICS BY INCOME QUINTILE

Variables	TOTAL (N=64,303)	Quintile 1 (N=13,060)	Quintile 2 (N=13,502)	Quintile 3 (N=12,808)	Quintile 4 (N=12,437)	Quintile 5 (N=12,496)	P-value
Age at diagnosis							
<=50	4,394 (6.8)	840 (6.4)	837 (6.2)	869 (6.8)	932 (7.5)	916 (7.3)	<.001
51-60	9,018 (14.0)	1,767 (13.5)	1,714 (12.7)	1,849 (14.4)	1,816 (14.6)	1,872 (15.0)	
61-70	15,436 (24.0)	2,963 (22.7)	3,195 (23.7)	3,095 (24.2)	3,072 (24.7)	3,111 (24.9)	
71-80	18,746 (29.2)	3,867 (29.6)	4,061 (30.1)	3,738 (29.2)	3,615 (29.1)	3,465 (27.7)	
>80	16,709 (26.0)	3,623 (27.7)	3,695 (27.4)	3,257 (25.4)	3,002 (24.1)	3,132 (25.1)	
Sex							
Female	31,370 (48.8)	6,671 (51.1)	6,659 (49.3)	6,218 (48.5)	5,853 (47.1)	5,969 (47.8)	<.001
Male	32,933 (51.2)	6,389 (48.9)	6,843 (50.7)	6,590 (51.5)	6,584 (52.9)	6,527 (52.2)	
Rural residence							
RIO <45	59,632 (92.7)	11,903 (91.1)	12,421 (92.0)	11,881 (92.8)	11,679 (93.9)	11,748 (94.0)	<.001
RIO =>45	4,671 (7.3)	1,157 (8.9)	1,081 (8.0)	927 (7.2)	758 (6.1)	748 (6.0)	
Elixhauser comorbidities							
<4	55,435 (86.2)	10,913 (83.6)	11,511 (85.3)	11,012 (86.0)	10,960 (88.1)	11,039 (88.3)	<.001
=>4	8,868 (13.8)	2,147 (16.4)	1,991 (14.7)	1,796 (14.0)	1,477 (11.9)	1,457 (11.7)	

Histology							
Other	1,858 (2.9)	348 (2.7)	382 (2.8)	375 (2.9)	357 (2.9)	396 (3.2)	0.192
Adenocarcinoma	62,445 (97.1)	12,712 (97.3)	13,120 (97.2)	12,433 (97.1)	12,080 (97.1)	12,100 (96.8)	
Stage at diagnosis							
I	12,126 (18.9)	2,268 (17.4)	2,517 (18.6)	2,460 (19.2)	2,416 (19.4)	2,465 (19.7)	<.001
II	16,062 (25.0)	3,309 (25.3)	3,404 (25.2)	3,158 (24.7)	3,064 (24.6)	3,127 (25.0)	
III	15,513 (24.1)	3,194 (24.5)	3,230 (23.9)	3,077 (24.0)	3,020 (24.3)	2,992 (23.9)	
IV	11,193 (17.4)	2,290 (17.5)	2,298 (17.0)	2,232 (17.4)	2,198 (17.7)	2,175 (17.4)	
unknown/missing	9,409 (14.6)	1,999 (15.3)	2,053 (15.2)	1,881 (14.7)	1,739 (14.0)	1,737 (13.9)	

Diagnostic pathway description

There were 11,378 (17.7%) patients with an asymptomatic interval. Patients living in the lowest income quintile neighbourhoods were less likely to experience an asymptomatic pathway (17.4% vs 20.4% in the highest income quintile). The first encounter for asymptomatic pathways differed slightly by neighbourhood income quintile. Patients living in the lowest income quintile neighbourhoods were less likely to have a lower GI scope as their first encounter (22.6% in Q1 vs 26.2% in Q5, p=0.0028) and more likely to have a gFOBT (77.4% in Q1 vs 73.9% in Q5, p=0.0033) compared to individuals residing the highest income neighbourhoods.

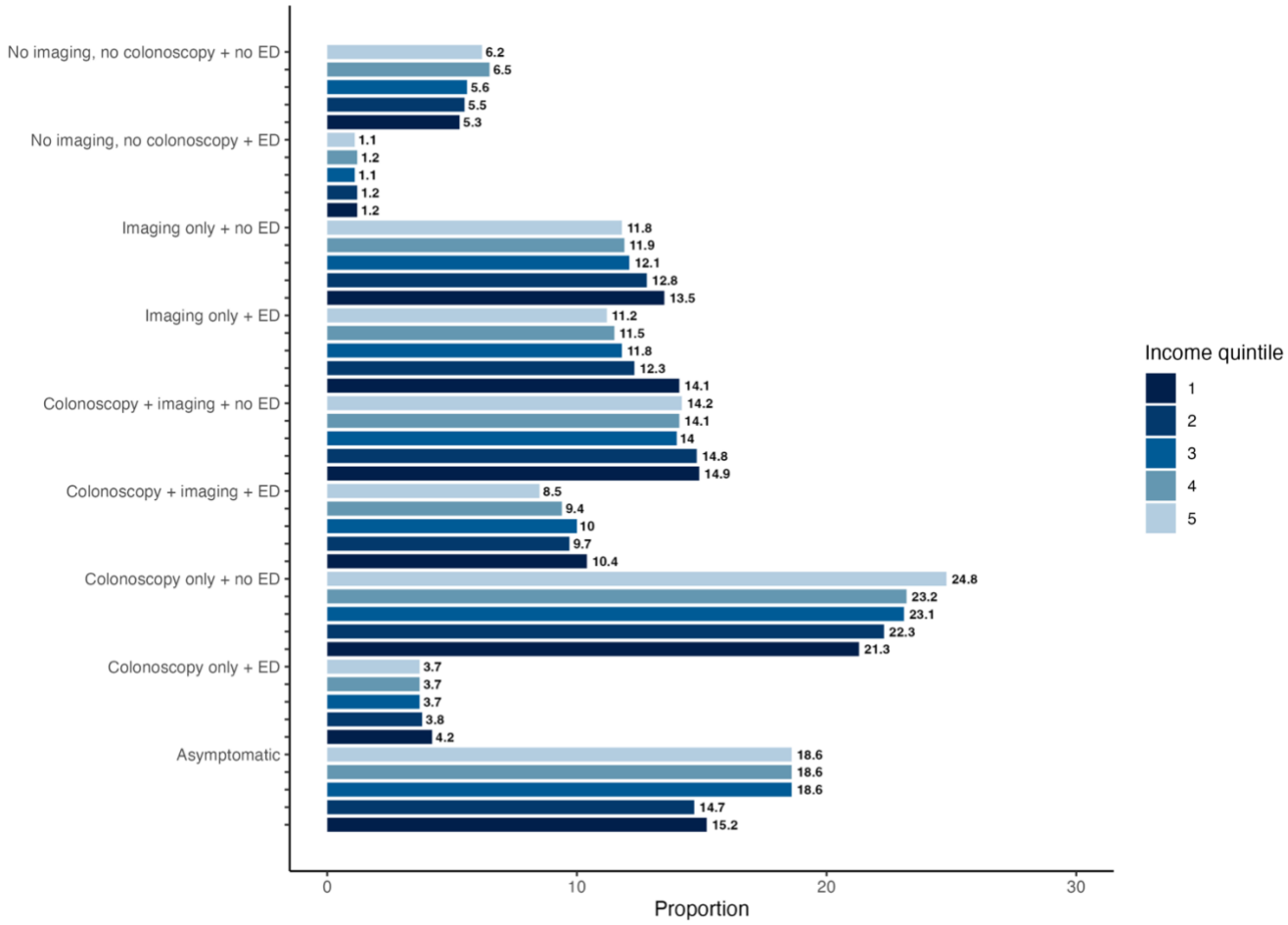
82,925 (82.3%) patients had a symptomatic interval. The first encounter for symptomatic pathways was most likely to be a diagnostic code for gastrointestinal signs and symptoms, followed by an emergency family physician visit. These were similar between neighbourhood income quintiles (Supplemental Table 3). An ED visit on the first encounter date was more likely to occur in patients living in the lowest income neighbourhoods (35.2%) compared to patients living in the highest income neighbourhoods (30.1%). Diagnostic pathways also differed by neighbourhood income quintile, with individuals residing in the lowest income neighbourhoods more likely to be diagnosed through lower GI endoscopy with imaging in the ED (12.2% vs

10.4% in quintile 5) and less likely to be diagnosed through colonoscopy only outside the ED (25.1% vs 30.5% in quintile 5) compared to individuals residing in the highest income neighbourhoods (Figure 2).

TABLE 6.2. FEATURES OF THE DIAGNOSTIC INTERVAL BY SYMPTOM STATUS AND INCOME QUINTILE (COLUMN PERCENT)

	ASYMPTOMATIC						SYMPTOMATIC					
	Quintile 1 (N=1,985)	**Quintile 2 (N=2,386)**	**Quintile 3 (N=2,377)**	**Quintile 4 (N=2,309)**	**Quintile 5 (N=2,321)**	**p-value**	**Quintile 1 (N=11,075)**	**Quintile 2 (N=11,116)**	**Quintile 3 (N=10,431)**	**Quintile 4 (N=10,128)**	**Quintile 5 (N=10,175)**	**p-value**
Diagnosed on index encounter date (diagnostic interval = 1 day)												
No	1,890 (95.2)	2,269 (95.1)	2,278 (95.8)	2,178 (94.3)	2,187 (94.2)	0.0702	10,141 (91.6)	10,254 (92.2)	9,634 (92.4)	9,335 (92.2)	9,320 (91.6)	0.0851
Yes	95 (4.8)	117 (4.9)	99 (4.2)	131 (5.7)	134 (5.8)		934 (8.4)	862 (7.8)	797 (7.6)	793 (7.8)	855 (8.4)	
Referring physician as first contact												
No	1,756 (88.5)	2,075 (87.0)	2,083 (87.6)	1,994 (86.4)	1,994 (85.9)	0.096	9,538 (86.1)	9,501 (85.5)	8,884 (85.2)	8,548 (84.4)	8,520 (83.7)	<.0001
Yes	229 (11.5)	311 (13.0)	294 (12.4)	315 (13.6)	327 (14.1)		1,537 (13.9)	1,615 (14.5)	1,547 (14.8)	1,580 (15.6)	1,655 (16.3)	
ED at index												
No	1,985 (100.0)	2,386 (100.0)	2,377 (100.0)	2,309 (100.0)	2,321 (100.0)	NA	7,177 (64.8)	7,476 (67.3)	7,023 (67.3)	6,926 (68.4)	7,117 (69.9)	<.0001
Yes							3,898 (35.2)	3,640 (32.7)	3,408 (32.7)	3,202 (31.6)	3,058 (30.1)	
Lower GI scope												
0	324 (16.3)	405 (17.0)	395 (16.6)	362 (15.7)	367 (15.8)	0.7301	4,451 (40.2)	4,286 (38.6)	3,923 (37.6)	3,866 (38.2)	3,788 (37.2)	<.0001
1+	1,661 (83.7)	1,981 (83.0)	1,982 (83.4)	1,947 (84.3)	1,954 (84.2)		6,624 (59.8)	6,830 (61.4)	6,508 (62.4)	6,262 (61.8)	6,387 (62.8)	

FIGURE 6.2 DIAGNOSTIC PATHWAYS BY INCOME QUINTILE



Diagnostic interval description

The diagnostic interval overall was 108 days (IQR 31-243 days) with a 90th percentile of 383 days. Patients with asymptomatic pathways had shorter median and 90th percentile diagnostic intervals compared to symptomatic pathways (median 71 days (IQR 35-137, 90th percentile 230 days) vs 148 median 121 days (IQR 29-273, 90th percentile 404 days), respectively). Asymptomatic median and 90th percentile diagnostic intervals were similar across neighbourhood income quintiles in descriptive analysis, ranging from 71 days (IQR 36-130, 90th percentile 222 days) among individuals residing in the lowest income neighbourhoods to 70 days (IQR 33-144, 90th percentile 228 days) in the highest income neighbourhoods (Supplemental

Table 4). Symptomatic diagnostic intervals ranged from 126 days (IQR 31-280, 90th percentile 410 days) among individuals residing in the lowest income neighbourhoods to 118 days (IQR 28-267, 90th percentile 400 days) in the highest income neighbourhoods (Table 3).

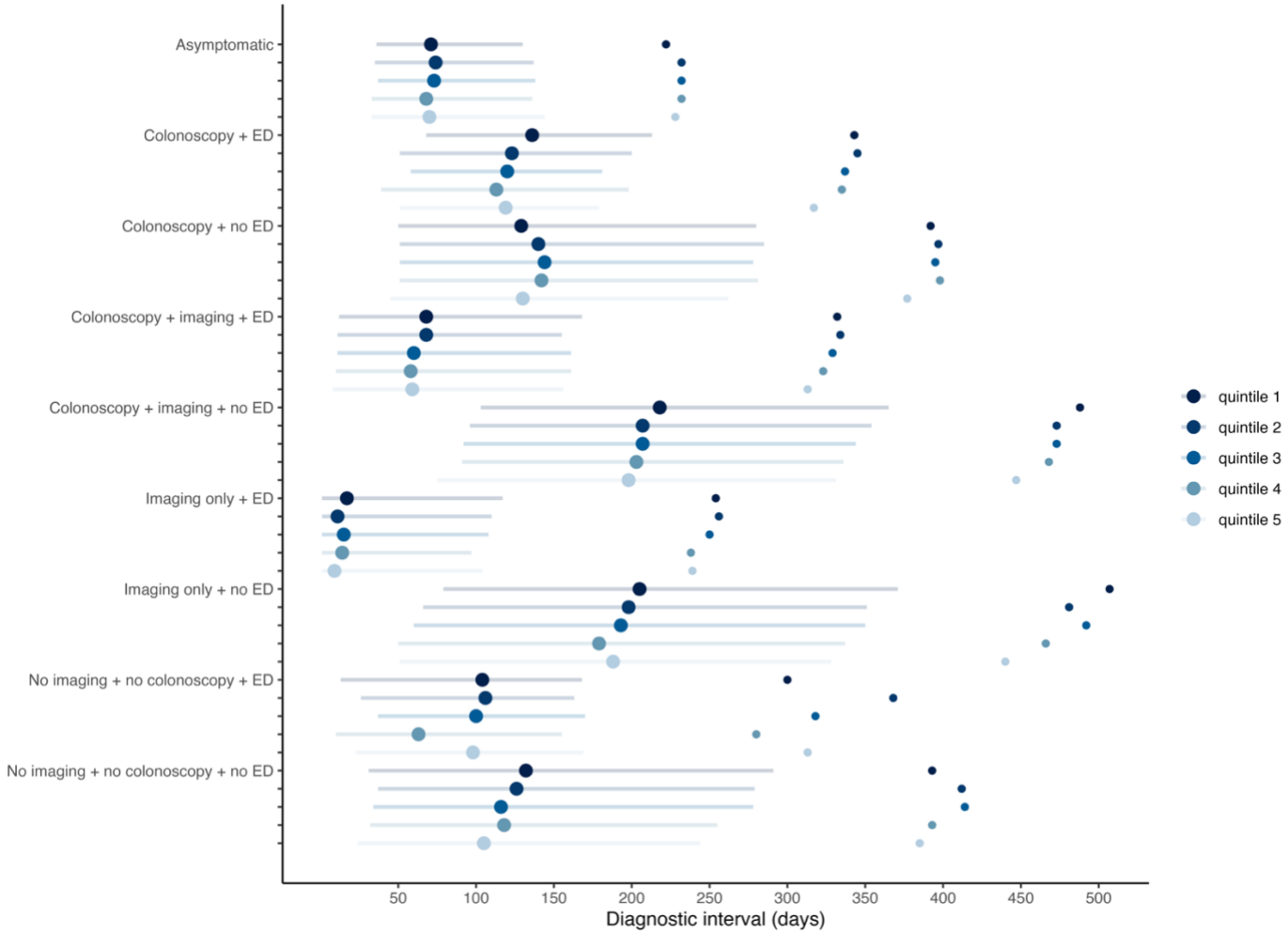
Older patients, women, those with comorbidities, or earlier stages had longer diagnostic intervals in both asymptomatic and symptomatic pathways. The nine diagnostic pathways had different diagnostic intervals, and these differed by income quintile, with individuals with the lowest income quintile generally experiencing longer diagnostic intervals across symptomatic pathways and similar intervals for the asymptomatic pathway compared to people with the highest income quintile (Figure 3).

TABLE 6.3. MEDIAN AND 90TH PERCENTILE DIAGNOSTIC INTERVAL BY SYMPTOM STATUS AND PATIENT AND DISEASE CHARACTERISTICS (SYMPTOMATIC PATHWAYS ONLY)

	SYMPTOMATIC									
	Q1		Q2		Q3		Q4		Q5	
Variable	Median (IQR)	90th pct	Median (IQR)	90th pct	Median (IQR)	90th pct	Median (IQR)	90th pct	Median (IQR)	90th pct
Diagnostic interval overall	126 (31-280)	410	124 (31-279)	410	122 (30-273)	406	118 (28-267)	400	116 (27-260)	392
Stage at diagnosis										
Stage I	160 (72-294)	408	153 (61-291)	407	148 (60-286)	396	153 (26-284)	405	153 (57-283)	400
Stage II	121 (30-277)	412	125 (35-270)	413	124 (35-266)	404	115 (29-261)	392	116 (30-256)	388
Stage III	126 (33-284)	405	119 (30-274)	400	116 (27-272)	407	113 (28-258)	402	113 (29-248)	383
Stage IV	85 (13-214)	382	82 (15-226)	385	79 (13-228)	388	70 (12-227)	376	59 (11-205)	360
Stage unknown/missing	145 (36-316)	435	138 (36-314)	434	139 (35-315)	434	140 (40-303)	423	137 (34-294)	428
Age at index (categorical)										
<=50	105 (18-246)	386	103 (22-229)	380	102 (20-258)	386	75 (17-207)	373	97 (22-211)	346
51-60	103 (19-233)	371	100 (22-240)	372	96 (17-214)	364	94 (19-226)	346	93 (19-206)	346
61-70	121 (29-265)	392	113 (29-248)	383	111 (29-254)	386	113 (25-255)	389	112 (27-257)	387
71-80	129 (35-286)	408	127 (35-284)	416	128 (38-282)	407	124 (31-272)	403	121 (31-260)	388
>80	147 (40-304)	447	144 (39-309)	445	147 (38-311)	442	142 (39-299)	434	141 (35-293)	425
Sex										
Female	137 (39-291)	418	135 (35-290)	422	128 (31-283)	413	126 (31-277)	408	120 (29-264)	395
Male	117 (25-266)	399	113 (29-261)	394	116 (30-263)	398	110 (26-256)	392	113 (26-254)	388
RIO at index										
Urban <45	128 (31-284)	413	125 (32-281)	411	123 (30-275)	407	119 (28-270)	400	117 (28-261)	395
Rural =>45	114 (28-228)	376	116 (28-255)	391	106 (29-249)	392	110 (26-245)	399	106 (22-235)	343
Elixhauser										
<4	119 (29-264)	398	116 (30-262)	398	114 (29-260)	396	112 (26-255)	393	110 (25-247)	382
=>4	173 (48-331)	466	171 (47-338)	467	164 (50-332)	464	168 (42-321)	451	170 (50-321)	438
Histology										
Other	113 (29-265)	416	113 (22-287)	423	112 (22-256)	399	113 (23-273)	410	111 (16-285)	433
Adenocarcinoma	127 (31-280)	409	124 (32-279)	409	122 (31-274)	407	118 (28-267)	400	116 (28-256)	390

Received at least one lower GI scope										
0	102 (10-266)	408	98 (11-257)	407	97 (11-255)	408	84 (10-238)	395	92 (8-240)	394
1+	141 (49-287)	411	136 (49-286)	411	136 (45-281)	404	136 (43-281)	402	130 (42-268)	390

FIGURE 6.3 MEDIAN AND INTERQUARTILE RANGE DIAGNOSTIC INTERVAL FOR DIAGNOSTIC PATHWAY BY INCOME QUINTILE (DOTS ON THE FAR RIGHT REPRESENT THE 90TH PERCENTILE, DOTS WITHIN THE LINES REPRESENT THE 50TH PERCENTILE, WITH THE END OF THE LEFT SIDE BEING THE 25TH PERCENTILE AND THE END OF THE RIGHT SIDE THE 75TH PERCENTILE)



Quantile Regression Models

For asymptomatic pathways, income was only significantly associated with the diagnostic interval at the 50th percentile, with patients in the three lowest income quintiles experiencing longer diagnostic intervals compared to patients in the highest income quintile. For symptomatic pathways, the three lowest income quintiles were associated with a longer 50th and 90th percentile diagnostic interval compared to patients in the highest income quintile (Table 4). For example, the 90th percentile diagnostic interval was 15 days (95% CI 6-23) longer for

patients in the lowest income quintile compared to the highest. After stratifying by stage, having low income was significantly associated with the diagnostic interval for asymptomatic patients with unknown or missing stage and symptomatic patients at stages 3 and 4 (Supplemental Table 5). For example, having the lowest income for symptomatic stage 3 patients was associated with a diagnostic interval that was 23 days (95% CI 8-38) longer compared to patients with the highest income.

TABLE 6.4. QUANTILE REGRESSION FOR THE EFFECT OF INCOME ON THE DIAGNOSTIC INTERVAL STRATIFIED BY SYMPTOM STATUS (REFERENCE = QUINTILE 5, ESTIMATES ARE IN DAYS).

	UNADJUSTED				ADJUSTED			
	50th percentile		90th percentile		50th percentile		90th percentile	
Model	Estimate (95% CI)	Wald p-value	Estimate (95% CI)	Wald p-value	Estimate (95% CI)	Wald p-value	Estimate (95% CI)	Wald p-value
ASYMPTOMATIC								
Intercept	70.00 (65.86-74.14)	<.0001	228.00 (216.13-239.87)	<.0001	48.77 (39.85-57.68)	<.0001	144.23 (130.48-157.97)	<.0001
Quintile 1	1.00 (-4.38-6.38)	0.2825	-6.00 (-26.70-14.70)	0.8742	4.83 (0.27-9.39)	0.0089	-7.96 (-15.92-(-0.01)	0.3745
Quintile 2	4.00 (-2.13-10.13)		4.00 (-15.88-23.88)		5.08 (-0.09-10.26)		-4.65 (-12.50-3.19)	
Quintile 3	3.00 (-2.43-8.43)		4.00 (-12.91-20.91)		7.22 (2.51-11.93)		-2.64 (-10.47-5.19)	
Quintile 4	-2.00 (-7.85-3.85)		4.00 (-15.25-23.25)		0.72 (-4.68-6.12)		-3.15 (-10.15-3.86)	
SYMPTOMATIC								
Intercept	116.00 (111.80-120.20)	<.0001	392.00 (385.61-398.39)	<.0001	108.48 (99.68-117.27)	<.0001	410.26 (395.75-424.77)	<.0001
Quintile 1	10.00 (4.42-15.58)	0.0046	18.00 (9.64-26.36)	<.0001	10.04 (4.37-15.71)	0.0056	14.76 (6.30-23.23)	0.0051
Quintile 2	8.00 (1.63-14.38)		18.00 (9.42-26.58)		5.73 (0.40-11.06)		11.53 (3.94-19.12)	
Quintile 3	6.00 (0.52-11.48)		14.00 (5.27-22.73)		5.65 (0.12-11.19)		9.55 (2.48-16.61)	
Quintile 4	2.00 (-4.44-8.44)		8.00 (-1.02-17.02)		1.08 (-4.45-6.60)		6.80 (-0.87-14.47)	

*Abbreviations : CI = confidence interval

DISCUSSION

This study found significantly longer symptomatic diagnostic intervals for patients experiencing the lowest income compared to those with the highest income, with increasing disparities with increasing stage at diagnosis. Our study found that the median and 90th percentile diagnostic interval for symptomatic pathways was 10 and 15 days longer for patients with the lowest income compared to those with the highest income. Smaller or no differences were found in the diagnostic interval by income for patients with asymptomatic pathways. Other studies examining inequalities in the diagnostic interval have found longer diagnostic intervals for patients residing in rural areas, women and immigrants.[34–36] These studies demonstrated a median interval that ranged from 18 days longer for rural patients to 5 days longer for new immigrants.[34–36] One other study demonstrated a median diagnostic interval of 6.5 days longer in patients living in low-income neighbourhoods compared to high-income neighbourhoods.[16] However, this study did not perform a multivariable analysis controlling for confounders.

Regardless of income, we found that patients with asymptomatic pathways had much shorter diagnostic intervals compared to symptomatic pathways, but patients with the lowest income were less likely to have asymptomatic diagnostic pathways and more likely to have pathways that included presenting in the ED compared to patients with the highest income. Asymptomatic pathways in our study reflect screening status as demonstrated by the similarity between the proportion of patients with asymptomatic pathways (18%) and previous studies indicating screening rates of 17% in Ontario.[29,37] This finding highlights similar known inequities in screening rates by income, gender and immigration status.[38]

Implications and future directions

Given the critical importance of timely diagnosis and treatment for outcomes such as patient anxiety and stage at diagnosis, the income-based differences we found in our study contribute to significant inequities within the cancer system. Delays in diagnosis for patients experiencing low income may compound with disparities in wait times across other aspects of the cancer care continuum, such as for treatment, and result in worse overall outcomes. For instance, a meta-analysis has shown that even a four-week delay between surgery and adjuvant treatment for colon cancer patients could increase the risk of mortality by 9-13%.[3] Such delays, coupled with well-known disparities in receiving cancer treatment, lead to substantial differences in cancer outcomes by income.[10]

The cancer system, especially in the diagnostic phase, is complex, fragmented, and often unfamiliar to patients. Well-educated and affluent patients may be able to absorb complex medical information, advocate for themselves, and have access to formal and informal healthcare networks, which are largely inaccessible to individuals experiencing low income or other structural disadvantages.[39] Population based screening programs, rapid assessment programs and patient navigation can improve access to cancer care and aims to reduce the time to diagnosis, especially for vulnerable populations.[40] Future research should assess these programs' effectiveness within subgroups of underserved individuals. Additionally, to implement and improve interventions aimed at shortening the diagnostic interval and improving outcomes, research should examine the pathways through which the diagnostic interval can be reduced and how these pathways may differ among structurally disadvantaged groups. For example, continuity of care with a regular family physician could facilitate screening and increase the likelihood of reporting signs and symptoms earlier, potentially resulting in shorter diagnostic

intervals for all patients, but may be especially important for patients experiencing low-income[41,42]

Strengths and limitations

This study has several strengths. We used routinely collected administrative data in Ontario that captures almost all cancer patients in the province; therefore, our results reflect real-world inequalities occurring in Ontarian colon cancer patients. We also used a modified definition of the diagnostic interval, allowing for more extended lookback periods to capture cancer and non-cancer-related diagnoses and procedures. This method might more likely capture intervals in patients experiencing diagnostic pathways that deviate from guideline recommendations. Finally, we used a conceptual model to determine the association between income and the diagnostic interval, which does not control for causal pathway variables.[33]

This study has limitations. Most importantly, individual income or other individual socioeconomic measures could not be obtained. While neighbourhood income may represent a measure of the neighbourhood environment, it is impossible to estimate the place-based effects of neighbourhood income without including individual income in a multi-level model.[43] In the absence of individual data, we used neighbourhood income to approximate individual income.[20,21] Studies estimating misclassification of individual income using neighborhood measures have demonstrated an attenuation of the effect of income on health outcomes when using neighbourhood income instead of individual income.[45,46] Therefore, we hypothesize that our results may underestimate the disparities in the diagnostic interval by income.[45] This limitation further stresses the importance of linking individual socioeconomic variables to rich, routinely collected administrative datasets. Second, we were unable to confirm screening status in the administrative data and therefore had to approximate screening with asymptomatic

pathways. It is possible that individuals may have received a colonoscopy for reasons other than screening, which might overestimate the number of individuals screened in our study. However, since the screening rates in our study were similar to those in the literature, we assume this misclassification is small. Finally, while our method for creating the diagnostic interval has been outlined in detail and used previously, it has yet to be validated due to limited access to detailed linked data.[24,25]

Conclusion

We found a meaningful differences in the diagnostic interval and pathways, with patients experiencing the lowest income less likely to be diagnosed through asymptomatic pathways, more likely to be diagnosed in the ED and having longer symptomatic diagnostic intervals compared to their high-income counterparts. Future work should examine inequalities in the diagnostic interval by individual income and among other vulnerable groups and determine pathways to reducing inequalities along the diagnostic interval, such as through improved access to screening programs, diagnostic navigation programs or regular contact with a family physician.

Author Statement

LE Davis: Conceptualization, methodology, formal analysis, data curation, writing – original draft, visualization, funding acquisition. **EC Strumpf:** Conceptualization, methodology, writing – review and editing, supervision, funding acquisition. **SV Patel:** Methodology, writing – review and editing. **AL Mahar:** Conceptualization, methodology, writing – review and editing, supervision, funding acquisition.

Funding: This study was supported by grants from the Canadian Institutes of Health Research (CIHR), the Canadian Cancer Society (MEGAN-CAN), the Canadian Centre for Applied Research in Cancer Control (ARCC) and Fonds de recherche du Québec – Santé (FRQS).

References

1. Brenner H, Jansen L, Ulrich A, Chang-Claude J, Hoffmeister M. Survival of patients with symptom- and screening-detected colorectal cancer. Oncotarget. 2016 May 17;7(28):44695–704.

2. Bretthauer M, Løberg M, Wieszczy P, Kalager M, Emilsson L, Garborg K, et al. Effect of Colonoscopy Screening on Risks of Colorectal Cancer and Related Death. N Engl J Med. 2022 Oct 27;387(17):1547–56.

3. Hanna TP, King WD, Thibodeau S, Jalink M, Paulin GA, Harvey-Jones E, et al. Mortality due to cancer treatment delay: systematic review and meta-analysis. BMJ. 2020 Nov 4;371:m4087.

4. Redaniel MT, Martin RM, Ridd MJ, Wade J, Jeffreys M. Diagnostic Intervals and Its Association with Breast, Prostate, Lung and Colorectal Cancer Survival in England: Historical Cohort Study Using the Clinical Practice Research Datalink. PLOS ONE. 2015 May 1;10(5):e0126608.

5. Tørring ML, Murchie P, Hamilton W, Vedsted P, Esteva M, Lautrup M, et al. Evidence of advanced stage colorectal cancer with longer diagnostic intervals: a pooled analysis of seven primary care cohorts comprising 11 720 patients in five countries. Br J Cancer. 2017 Sep;117(6):888–97.

6. Weller D, Vedsted P, Rubin G, Walter FM, Emery J, Scott S, et al. The Aarhus statement: improving design and reporting of studies on early cancer diagnosis. Br J Cancer. 2012 Mar;106(7):1262–7.

7. Din NU, Ukoumunne OC, Rubin G, Hamilton W, Carter B, Stapley S, et al. Age and Gender Variations in Cancer Diagnostic Intervals in 15 Cancers: Analysis of Data from the UK Clinical Practice Research Datalink. PLOS ONE. 2015 May 15;10(5):e0127717.

8. Lambert LK, Horrill TC, Beck SM, Bourgeois A, Browne AJ, Cheng S, et al. Health and healthcare equity within the Canadian cancer care sector: a rapid scoping review. Int J Equity Health. 2023 Jan 28;22(1):20.

9. International Agency for Research on Cancer. Reducing social inequalities in cancer: evidence and priorities for research. Lyon, France: World Health Organization; 2019 Apr.

10. Davis LE, Coburn NG, Hallet J, Earle CC, Liu Y, Myrehaug S, et al. Material deprivation and access to cancer care in a universal health care system. Cancer. 2020;126(20):4545–52.

11. Konradsen AA, Lund CM, Vistisen KK, Albieri V, Dalton SO, Nielsen DL. The influence of socioeconomic position on adjuvant treatment of stage III colon cancer: a systematic review and meta-analysis. Acta Oncol. 2020 Nov 1;59(11):1291–9.

12. Coughlin SS. Social determinants of colorectal cancer risk, stage, and survival: a systematic review. Int J Colorectal Dis. 2020 Jun 1;35(6):985–95.

13. Golding H, Webber CE, Groome PA. Factors contributing to time to diagnosis in symptomatic colorectal cancer: A scoping review. Eur J Cancer Care (Engl). 2020 Dec;30:e13397.

14. Walter FM, Rubin G, Bankhead C, Morris HC, Hall N, Mills K, et al. Symptoms and other factors associated with time to diagnosis and stage of lung cancer: a prospective cohort study. Br J Cancer. 2015 Mar;112(1):S6–13.

15. Webber C, Jiang L, Grunfeld E, Groome PA. Identifying predictors of delayed diagnoses in symptomatic breast cancer: a scoping review. Eur J Cancer Care (Engl). 2017;26(2):e12483.

16. Sikdar KC, Dickinson J, Winget M. Factors associated with mode of colorectal cancer detection and time to diagnosis: a population level study. BMC Health Serv Res. 2017 Dec;17(1):1–11.

17. Clarke E, Marrett L, Kreiger N. Cancer registration in Ontario: A computer approach. In: Cancer registration principles and methods. Lyon, France: IARC Publication; 1991. p. 246–57.

18. Juurlink D, Preyra C, Croxford R, Chong A, Austin P, Tu J, et al. Canadian institute for health information discharge abstract database: a validation study. ICES investigative report. Toronto: Institute for Clinical Evaluative Sciences; 2006.

19. Gibson D, Richards H, Chapman A. The National Ambulatory Care Reporting System: factors that affect the quality of its emergency data. Int J Inf Qual. 2008 Jan;2(2):97–114.

20. CIHI. Trends in Income-Related Health Inequalities in Canada [Internet]. Ottawa, Ontario: Canadian Institutes of Health Information; 2015. Available from: https://secure.cihi.ca/free_products/trends_in_income_related_inequalities_in_canada_2015_en.pdf

21. Booth CM, Li G, Zhang-Salomons J, Mackillop WJ. The impact of socioeconomic status on stage of cancer at diagnosis and survival: a population-based study in Ontario, Canada. Cancer. 2010 Sep 1;116:4160–7.

22. R Wilkins. PCCF + Version 3G Users Guide: Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files. Cat. No. 82F0086-XDB. 2001.

23. Dissemination area: Detailed definition [Internet]. [cited 2022 Jun 2]. Available from: https://www150.statcan.gc.ca/n1/pub/92-195-x/2011001/geo/da-ad/def-eng.htm

24. Groome PA, Webber C, Whitehead M, Moineddin R, Grunfeld E, Eisen A, et al. Determining the Cancer Diagnostic Interval Using Administrative Health Care Data in a Breast Cancer Cohort. JCO Clin Cancer Inform. 2019 Dec 1;(3):1–10.

25. Webber C, Flemming JA, Birtwhistle R, Rosenberg M, Groome PA. Colonoscopy resource availability and its association with the colorectal cancer diagnostic interval: A population-based cross-sectional study. Eur J Cancer Care (Engl). 2020;29(1):e13187.

26. Webber C. Availability and quality of colonoscopy resource and the colorectal cancer diagnostic interval [Internet] [PhD Thesis]. [Kingston, ON]: Queen's University; 2017. Available from: https://qspace.library.queensu.ca/handle/1974/6707

27. Castelo M, Paszat L, Hansen BE, Scheer AS, Faught N, Nguyen L, et al. Measurement of clinical delay intervals among younger adults with colorectal cancer using health administrative data: a population-based analysis. BMJ Open Gastroenterol. 2022 Nov 1;9(1):e001022.

28. Benneyan J, Lloyd R, Plsek P. Statistical process control as a tool for research and healthcare improvement. Qual Saf Health Care. 2003 Dec;12(6):458–64.

29. Guan Z, Webber C, Flemming JA, Mavor ME, Whitehead M, Chen BE, et al. Real-world colorectal cancer diagnostic pathways in Ontario, Canada: A population-based study. Eur J Cancer Care (Engl). 2022;31(5):e13603.

30. Lieffers JR, Baracos VE, Winget M, Fassbender K. A comparison of Charlson and Elixhauser comorbidity measures to predict colorectal cancer survival using administrative health data. Cancer. 2011 May 1;117(9):1957–65.

31. Kralj B. Measuring 'rurality' for purposes of health-care planning: an empirical measure for Ontario. Ont Med Rev. 2000;33–52.

32. Hao L, Naiman DQ. Quantile Regression. SAGE; 2007. 142 p.

33. Jackson JW. Meaningful causal decompositions in health equity research: definition, identification, and estimation through a weighting framework. Epidemiol Camb Mass. 2021 Mar 1;32(2):282–90.

34. Cohn JA, Vekhter B, Lyttle C, Steinberg GD, Large MC. Sex disparities in diagnosis of bladder cancer after initial presentation with hematuria: A nationwide claims-based investigation. Cancer. 2014;120(4):555–61.

35. Lofters AK, McBride ML, Li D, Whitehead M, Moineddin R, Jiang L, et al. Disparities in breast cancer diagnosis for immigrant women in Ontario and BC: results from the CanIMPACT study. BMC Cancer. 2019 Jan 9;19(1):42.

36. Bergin RJ, Emery J, Bollard RC, Falborg AZ, Jensen H, Weller D, et al. Rural-Urban Disparities in Time to Diagnosis and Treatment for Colorectal and Breast Cancer. Cancer Epidemiol Biomark Prev. 2018;27(9):1036–46.

37. Weller D, Menon U, Falborg AZ, Jensen H, Barisic A, Knudsen AK, et al. Diagnostic routes and time intervals for patients with colorectal cancer in 10 international jurisdictions;

findings from a cross-sectional study from the International Cancer Benchmarking Partnership (ICBP). BMJ Open. 2018 Nov 1;8(11):e023870.

38. Mosquera I, Mendizabal N, Martín U, Bacigalupe A, Aldasoro E, Portillo I, et al. Inequalities in participation in colorectal cancer screening programmes: a systematic review. Eur J Public Health. 2020 Jun 1;30(3):558–67.

39. Dixit N, Rugo H, Burke NJ. Navigating a Path to Equity in Cancer Care: The Role of Patient Navigation. Am Soc Clin Oncol Educ Book. 2021 Jun;(41):3–10.

40. Gagliardi A, Grunfeld E, Evans WK. Evaluation of Diagnostic Assessment Units in Oncology: A Systematic Review. J Clin Oncol. 2004 Mar 15;22(6):1126–35.

41. Gupta S, Brenner AT, Ratanawongsa N, Inadomi JM. Patient Trust in Physician Influences Colorectal Cancer Screening in Low-Income Patients. Am J Prev Med. 2014 Oct 1;47(4):417–23.

42. Lofters A, Salahub C, Austin PC, Bai L, Berthelot S, Bhatia RS, et al. Up-to-date on cancer screening among Ontario patients seen by walk-in clinic physicians: A retrospective cohort study. Prev Med. 2023 Jul;172:107537.

43. Merlo J, Chaix B, Yang M, Lynch J, Råstam L. A brief conceptual tutorial on multilevel analysis in social epidemiology: interpreting neighbourhood differences and the effect of neighbourhood characteristics on individual health. J Epidemiol Community Health. 2005 Dec 1;59(12):1022–9.

44. Lupton R. Neighbourhood Effects: Can We Measure Them and Does it Matter? [Internet]. Rochester, NY; 2003 [cited 2022 Oct 21]. Available from: https://papers.ssrn.com/abstract=1158964

45. Davis LE, Mahar AL, Strumpf EC. Agreement between individual and neighborhood income measures in patients with colorectal cancer in Canada. JNCI J Natl Cancer Inst. 2023 Jan 27;djad017.

46. Buajitti E, Chiodo S, Rosella LC. Agreement between area- and individual-level income measures in a population-based cohort: Implications for population health research. SSM - Popul Health. 2020 Apr 1;10:100553.

6.3 Supplemental material

Supplemental Table 1. Data sources

Database	Description
Ontario Cancer Registry (OCR)	The OCR is a passive provincial registry that captures over 95% of incident cancer diagnoses in Ontario since 1964. It includes information on primary cancer site, diagnosis dates, histology, and stage at diagnosis. We used the OCR to capture incident colon cancer diagnoses, date of diagnosis and stage at diagnosis.
Ontario Health Insurance Plan Database (OHIP)	OHIP contains billing claims made by all Ontario physicians, including inpatient and outpatient settings. Each claim includes the date, one fee code representing the billable service and one diagnosis code, physician specialty and referring physician where applicable. Physicians are required to submit a diagnosis with each fee code.
CIHI Discharge Abstract Database (DAD) and Same Day Surgery Database (SDS)	CIHI DAD and SDS are mandatory reporting systems that provide information on hospital admissions and same-day surgeries. Each record includes up to 20 intervention codes and 25 diagnosis codes.
CIHI National Ambulatory Care Reporting System (NACRS)	All emergency department visits in Ontario are captured in NACRS, including administrative, demographic, and clinical data. Each record includes up to 10 intervention codes and 10 diagnosis codes.
Registered Persons Database (RPDB)	The RPDB is an ICES database derived from all administrative data sources and provides basic demographic information, such as age, sex, postal code, date of last contact with the healthcare system and OHIP eligibility.

Supplemental Table 2. Covariate definitions

Variable	Source	Definition	Type
Socio-demographic characteristics			
Age	RPDB	Age at diagnosis	Categorical: ≤50, 51-60, 61-70, 71-80, >80. Continuous.
Sex	RPDB	Sex	Categorical: Male/Female
Rural residence	RPDB	Rural residence was measured by linking postal codes at the time of diagnosis to the Rurality Index of Ontario (RIO). RIO is a function of a region's population size as well as access to healthcare, such as distance to family practitioners and travel time.	Categorical: Rural (≥45)/Urban (<45)
Comorbidities	DAD, SDS, OHIP	Measured using the Elixhauser Comorbidity, which measures hospitalizations in the two years prior to cancer diagnosis.	Categorical: <4; =>4
Disease characteristics			
Histology	OCR	Histology was broadly categorized as adenocarcinoma or non-adenocarcinoma. Adenocarcinoma was defined using the following morphology codes: - 80003, 81403, 81406, 81443, 81483, 81563, 82013, 82103, 82113, 82133, 82403, 82413, 82433, 82443, 82453, 82463, 82493, 82503, 82533, 82553, 82603, 82613, 82623, 82633, 82653, 83103, 83123, 83233, 83373, 83413, 83803, 84803, 84813, 84903, 85743	Adenocarcinoma/non-adenocarcinoma
Stage at diagnosis	OCR	Stage at diagnosis represents best UICC/AJCC stage which is a combination of Collaborative Staging approach and data from medical records at regional cancer centres.	Categorical: I/II/III/IV/unknown or missing
Diagnosis year	OCR	Year of index colon cancer diagnosis.	Categorical: 2007-2019
Diagnostic interval characteristics			

Symptom status	OHIP, DAD, SDS, NACRS	Asymptomatic if the patients received a gFOBT or lower GI endoscopy as their first healthcare encounter and there were no encounters in the ED on the first encounter date. - Lower GI endoscopy: OHIP = 547, 548, 546, Z555, Z498, Z495, Z494, Z496, Z499, Z497, Z492, Z580, Z571, Z765, Z491, Z570, Z543, Z536, Z535. CIHI = 2NM70^^, 1NM87BA, 1NQ87BA^, 1NM87DA, 1NQ87DA, 2NQ70^^, 2NM70BABG, 2NM70BABH - gFOBT: OHIP = L181, L179, Q152, Q150, G004, Q005	Categorical: Symptomatic/Asymptomatic
ED at index	OHIP, NACRS	First encounter occurs in the ED and that record is present in NACRS or is with an emergency department physician. - ED physician visit: OHIP = H055, A888, H102, H103, H101, H104, H132, H131, H133, H134, H122, H121, H123, H124, H152, H151, H153, H154, H065, H105, C933, A933	Categorical: ED/non-ED
Lower GI scope	OHIP, DAD, SDS, NACRS	Measured at any time during the diagnostic interval (time from first healthcare encounter to diagnosis) - Lower GI endoscopy: OHIP = 547, 548, 546, Z555, Z498, Z495, Z494, Z496, Z499, Z497, Z492, Z580, Z571, Z765, Z491, Z570, Z543, Z536, Z535. CIHI = 2NM70^^, 1NM87BA, 1NQ87BA^, 1NM87DA, 1NQ87DA, 2NQ70^^, 2NM70BABG, 2NM70BABH	Categorical: 0/1+
Imaging	OHIP, DAD, SDS, NACRS	Measured at any time during the diagnostic interval. Colon-cancer related imaging if in any of the following: - Abdominal or pelvis CT: OHIP = X410, X126, X409, X234, X233,	Categorical: 0/1+

		X232, X231, X125, X407, X406. CIHI = 3NM20, 3OT20, 3GY20, 3ZZ20 - Abdominal or pelvis MRI: OHIP = X451, X455, X461, X465. CIHI = 3OT40 - Abdominal or pelvis ultrasound: OHIP = J435, J428, J162, J462, J138, J463, J163, J438, J135, J128. CIHI = 3OT30^^ - Abdominal or pelvis ultrasound: OHIP = X100, X101, X090, X092, X091, X039, X197, X195, X113, X112, X104, X103, X111, X036, X038, X037. CIHI = 3NL10, 3OT10, 3GY10, 3SL10, 3OT12, 3NQ10, 3NZ10, 3NM10, 3NA10, 3NK10, 3SQ10, 3VA10	
Diagnostic pathway summary		Summary variable using a combination of the following variables defined above: Symptom status, ED at index, lower GI scope and imagining.	Categorical: 1 = Asymptomatic; 2 = Lower GI scope + no imagining + ED at index; 3 = Lower GI scope + no imagining + no ED at index; 4 = Lower GI scope + imagining + ED at index; 5 = Lower GI scope + imagining + no ED at index; 6 = No lower GI scope + imagining + ED at index; 7 = No lower GI scope + imagining + no ED at index; 8 = No lower GI scope + no imagining + ED at index; 9 = No lower GI scope + no imagining + no ED at index;

Supplemental Table 3. Description of the first encounter for patients with symptomatic diagnostic pathways by neighbourhood income quintile (proportion of total quintile for each category, patients can have multiple first encounters)

Encounter Category	SYMPTOMATIC (ALL)									
	Q1 (N=10,509)		Q2 (N=10,496)		Q3 (N=9,808)		Q4 (N=9,483)		Q5 (N=9,471)	
	N	%	N	%	N	%	N	%	N	%
Colorectal cancer	360	3.43	315	3	312	3.18	313	3.3	352	3.72
Other cancer	344	3.27	367	3.5	338	3.45	339	3.57	334	3.53
Signs and symptoms Gi	4720	44.91	4775	45.49	4314	43.98	4323	45.59	4381	46.26
Signs and symptoms liver	23	0.22	46	0.44	30	0.31	39	0.41	38	0.4
Signs and symptoms haematology	478	4.55	528	5.03	458	4.67	444	4.68	431	4.55
Anemia	1417	13.48	1390	13.24	1326	13.52	1222	12.89	1109	11.71
Signs and symptoms nutritional	42	0.4	33	0.31	34	0.35	47	0.5	42	0.44
Signs and symptoms bacterial or viral	44	0.42	49	0.47	30	0.31	40	0.42	27	0.29
Abdominal or pelvis MRI	20	0.19	14	0.13	15	0.15	18	0.19	20	0.21
Biopsy	773	7.36	735	7	729	7.43	673	7.1	744	7.86
Abdominal ultrasound	215	2.05	237	2.26	213	2.17	214	2.26	215	2.27
Head or spine MRI	7	0.07	17	0.16	11	0.11	5	0.05	5	0.05
Emergency FP visit	2187	20.81	2115	20.15	1905	19.42	1886	19.89	1700	17.95
Critical care	437	4.16	378	3.6	343	3.5	302	3.18	299	3.16
Gastroenterologist consult	182	1.73	196	1.87	221	2.25	194	2.05	237	2.5
Lower GI scope	731	6.96	729	6.95	706	7.2	666	7.02	757	7.99
Upper GI scope	574	5.46	582	5.54	515	5.25	514	5.42	612	6.46
Non-GI scope	24	0.23	32	0.3	18	0.18	18	0.19	14	0.15
gFOBT	14	0.13	19	0.18	12	0.12	10	0.11	13	0.14
Colon resection	464	4.42	429	4.09	418	4.26	395	4.17	399	4.21
Abdominal xray	1528	14.54	1386	13.21	1313	13.39	1242	13.1	1143	12.07
Other abdominal procedure	140	1.33	131	1.25	124	1.26	108	1.14	109	1.15
Miscellaneous procedure	7	0.07	15	0.14	10	0.1	13	0.14	9	0.1
Abdominal CT	1754	16.69	1698	16.18	1565	15.96	1491	15.72	1420	14.99
Head or spine CT	362	3.44	365	3.48	328	3.34	308	3.25	285	3.01
General surgery consult	689	6.56	636	6.06	630	6.42	691	7.29	681	7.19
Cardiovascular visit	543	5.17	503	4.79	442	4.51	438	4.62	510	5.38
Other consultation	181	1.6	225	1.95	160	1.51	190	1.9	168	1.64

Supplemental Table 4. Median and 90th percentile diagnostic interval by symptom status and patient and disease characteristics (asymptomatic pathways only)

	ASYMPTOMATIC									
	Q1		Q2		Q3		Q4		Q5	
Variable	Median (IQR)	90th pct	Median (IQR)	90th pct	Median (IQR)	90th pct	Median (IQR)	90th pct	Median (IQR)	90th pct
Diagnostic interval overall	71 (36-130)	222	74 (35-137)	232	73 (37-138)	232	68 (33-136)	232	70 (33-144)	228
Stage at diagnosis										
Stage I	85 (41-141)	235	88 (41-142)	218	81 (44-148)	223	78 (35-141)	240	78 (36-148)	218
Stage II	69 (37-130)	218	66 (33-130)	230	65 (36-129)	222	66 (35-135)	220	68 (35-143)	231
Stage III	73 (37-128)	240	77 (36-142)	242	67 (35-135)	232	68 (37-129)	240	64 (31-132)	229
Stage IV	74 (30-148)	232	65 (30-145)	253	85 (37-159)	275	65 (31-143)	259	85 (35-169)	255
Stage unknown/missing	58 (29-113)	173	64 (31-125)	229	70 (35-126)	202	52 (22-127)	207	51 (28-1289)	205
Age at index (categorical)										
<=50	70 (29-120)	194	53 (20-115)	172	53 (22-107)	179	44 (18-85)	167	51 (16-123)	178
51-60	71 (34-125)	219	67 (34-136)	222	67 (33-135)	219	71 (31-140)	230	71 (31-134)	231
61-70	70 (38-125)	210	72 (33-133)	220	72 (39-131)	227	69 (34-128)	215	70 (33-142)	219
71-80	70 (33-137)	239	78 (37-140)	245	76 (38-148)	236	67 (35-148)	255	72 (36-150)	245
>80	79 (39-149)	241	80 (39-151)	250	88 (44-164)	254	73 (40-143)	251	69 (37-159)	235
Sex										
Female	75 (36-141)	218	72 (36-144)	242	72 (39-138)	245	66 (30-137)	230	73 (33-148)	234
Male	69 (36-125)	224	76 (34-133)	224	74 (36-138)	220	70 (36-135)	233	65 (33-140)	225
RIO at index										
Urban <45	71 (36-132)	219	75 (36-139)	227	73 (37-137)	232	68 (34-136)	231	70 (33-146)	231
Rural =>45	77 (43-125)	239	59 (24-124)	271	67 (39-149)	228	75 (33-141)	234	64 (30-110)	204
Elixhauser										
<4	71 (36-128)	215	72 (34-133)	226	72 (36-136)	227	67 (33-133)	226	68 (32-141)	226

=>4	93 (53-172)	288	121 (70-210)	294	91 (44-176)	268	137 (57-196)	276	123 (64-209)	268
Histology										
Other	85 (36-120)	244	58 (24-125)	188	98 (49-213)	324	44 (17-92)	164	49 (32-149)	234
Adenocarcinoma	71 (36-131)	222	74 (35-137)	233	72 (37-137)	227	69 (34-137)	232	70 (33-144)	228
Received at least one lower GI scope										
0	86 (47-157)	276	82 (38-148)	250	89 (46-166)	278	78 (36-160)	264	75 (37-156)	255
1+	70 (35-127)	212	72 (34-133)	225	71 (36-133)	217	66 (33-133)	222	68 (32-142)	224

Supplemental Table 5. Quantile regression for the effect of neighbourhood income quintile on the diagnostic interval. Stratified by symptom status and stage at diagnosis. Multivariable intercept represents the estimated diagnostic interval in days at baseline (for an individual with mean age 71, male, income quintile 5, stage unknown, diagnosed in 2019)

	UNADJUSTED				ADJUSTED			
	50th percentile		90th percentile		50th percentile		90th percentile	
Model	Estimate (95% CI)	Wald p-value	Estimate (95% CI)	Wald p-value	Estimate (95% CI)	Wald p-value	Estimate (95% CI)	Wald p-value
ASYMPTOMATIC								
Stage 1								
Intercept	78.00 (68.49-87.15)	<.0001	217.69 (201.25-234.13)	<.0001	54.78 (27.37-82.19)	<.0001	121.75 (62.34-181.15)	<.0001
Q1	7.00 (-6.37-20.37)	0.4027	17.31 (-17.59-52.20)	0.6643	7.32 (-3.75-18.40)	0.1938	-8.50 (-25.94-8.94)	0.5398
Q2	10.00 (-2.29-22.29)		0.31 (-26.96-27.58)		10.11 (0.23-19.98)		-8.25 (-23.74-7.24)	
Q3	3.00 (-8.04-14.04)		5.31 (-19.63-30.25)		10.11 (-0.18-20.40)		-3.00 (-18.91-12.91)	
Q4	0.00 (-13.84-13.84)		22.31 (-10.89-55.51)		2.04 (-8.87-12.95)		4.25 (-13.33-21.83)	
Stage 2								
Intercept	68.00 (61.02-74.98)	<.0001	231.00 (200.07-261.93)	<.0001	39.00 (18.98-59.01)	0.0001	134.16 (101.86-166.46)	<.0001
Q1	1.00 (-9.21-11.21)	0.8806	-13.00 (-59.97-33.97)	0.9624	4.09 (-5.20-13.38)	0.9102	-7.18 (-22.97-8.61)	0.3869
Q2	-2.00 (-11.46-7.46)		-1.00 (-64.86-62.86)		0.92 (-8.64-10.48)		-1.42 (-16.44-13.59)	
Q3	-3.00 (-11.39-5.39)		-9.00 (-49.59-31.59)		3.73 (-6.35-13.81)		-15.39 (-33.14-2.35)	
Q4	-2.00 (-11.17-7.17)		-11.00 (-50.17-28.17)		2.09 (-7.72-11.90)		-8.00 (-24.37-8.37)	
Stage 3								
Intercept	64.00 (57.02-70.98)	<.0001	229.00 (201.56-256.44)	<.0001	44.81 (21.10-68.52)	0.0002	167.32 (131.22-203.42)	<.0001
Q1	9.26 (-0.62-19.14)	0.0238	10.12 (-30.08-50.32)	0.9556	10.25 (0.49-20.00)	0.1299	-2.33 (-21.69-17.03)	0.8655
Q2	13.08 (4.37-21.78)		13.00 (-24.41-50.41)		12.00 (0.56-23.44)		7.53 (-9.90-24.96)	
Q3	3.00 (-7.62-13.62)		3.00 (-29.25-35.25)		5.43 (-4.21-15.08)		3.08 (-17.10-23.26)	
Q4	4.00 (-5.56-13.56)		11.00 (-32.14-54.14)		4.43 (-5.42-14.29)		1.79 (-16.58-20.16)	
Stage 4								
Intercept	85.00 (70.53-99.47)	<.0001	255.00 (224.12-285.88)	<.0001	108.91 (67.63-150.18)	<.0001	176.77 (127.88-225.65)	<.0001
Q1	-12.29 (-31.33-6.74)	0.0290	-23.00 (-72.15-26.15)	0.5181	-11.91 (32.47-8.65)	0.1876	-21.47 (-42.33-(-0.60))	0.0896
Q2	-20.00 (-41.02-1.02)		-2.00 (-58.96-54.96)		-15.62 (-31.59-0.35		-16.23 (-37.29-4.83)	

Q3	0.00 (-20.72-20.72)		20.00 (-25.45-65.45)		-8.26 (-27.54-11.03)		4.60 (-17.41-26.62)	
Q4	-19.58 (-36.19-(-2.96))		4.00 (-36.92-44.92)		-18.57 (-35.15-(-1.99)		-10.88 (-30.13-8.36)	
Stage unknown								
Intercept	51.32 (43.15-59.49)	<.0001	205.00 (163.76-246.24)	<.0001	45.59 (29.74-61.44)	<.0001	131.57 (99.61-163.53)	<.0001
Q1	6.68 (-2.78-16.14)	0.0130	-32.00 (-87.37-23.37)	0.1581	4.64 (-6.31-15.59)	0.0117	-26.00 (-47.22-(-4.78)	0.0277
Q2	12.68 (0.60-24.76)		24.00 (-23.71-71.71)		8.68 (-2.25-19.61)		-11.32 (-35.36-12.73)	
Q3	18.68 (6.33-31.03)		-3.00 (-64.42-58.42)		15.44 (3.58-27.30)		4.95 (-20.45-30.35)	
Q4	0.68 (-14.43-15.80)		2.00 (-55.09-59.09)		-4.20 (-15.71-7.31)		-11.47 (-36.84-13.89)	
SYMPTOMATIC								
Stage 1								
Intercept	153.00 (145.46-160.54)	<.0001	400.00 (387.80-412.20)	<.0001	113.52 (71.55-155.48)	<.0001	435.94 (374.47-497.41)	<.0001
Q1	7.00 (-3.39-17.39)	0.3191	8.00 (-8.57-24.57)	0.5740	6.44 (-6.42-19.30)	0.6874	5.60 (-9.66-20.86)	0.7480
Q2	0.00 (-11.37-11.37)		7.00 (-8.74-22.74)		0.86 (-11.86-13.58)		6.78 (-7.30-20.85)	
Q3	-5.00 (-15.90-5.90)		-4.00 (-23.22-15.22)		-3.79 (-16.47-8.90)		4.71 (-9.86-19.29)	
Q4	0.17 (-12.48-12.83)		5.00 (-12.53-22.53)		-0.15 (-13.02-12.71)		10.31 (-4.76-25.38)	
Stage 2								
Intercept	116.00 (108.86-123.14)	<.0001	388.00 (373.95-402.05)	<.0001	98.09 (69.18-127.01)	<.0001	395.89 (357.45-434.33)	<.0001
Q1	5.00 (-4.99-14.99)	0.1636	24.00 (5.76-42.24)	0.0089	5.85 (-6.54-18.25)	0.3688	14.67 (0.47-28.88)	0.2024
Q2	9.00 (-0.29-18.29)		25.00 (6.90-43.10)		5.47 (-5.02-15.96)		7.06 (6.72-20.83)	
Q3	8.00 (-3.05-19.05)		16.00 (-2.05-34.05)		8.46 (-2.09-19.00)		5.60 (-10.16-21.37)	
Q4	-1.00 (-11.33-9.33)		4.00 (-14.19-22.19)		-0.95 (-11.71-9.81)		-1.28 (-15.79-13.22)	
Stage 3								
Intercept	113.00 (105.26-120.74)	<.0001	383.00 (371.80-394.20)	<.0001	114.29 (93.62-134.96)	<.0001	423.14 (385.04-461.25)	<.0001
Q1	13.00 (1.43-24.57)	0.1417	22.00 (6.11-37.89)	0.0206	11.78 (2.46-21.09)	0.0552	23.09 (7.74-38.44)	0.0515
Q2	6.00 (-5.75-17.75)		17.00 (2.28-31.73)		2.83 (-8.92-14.58)		12.65 (-3.51-28.80)	
Q3	3.00 (-6.84-12.84)		24.00 (7.02-40.98)		1.78 (-8.70-12.25)		16.41 (0.40-32.42)	
Q4	0.00 (-9.79-9.79)		19.00 (2.37-35.63)		-1.62 (-13.00-9.76)		18.32 (2.32-34.32)	
Stage 4								
Intercept	59.00 (49.07-68.93)	<.0001	360.00 (340.20-379.80)	<.0001	62.25 (40.42-84.08)	<.0001	373.61 (332.53-414.69)	<.0001

Q1	26.00 (8.02-43.98)	0.0164	22.00 (-2.76-46.76)	0.1724	14.70 (2.91-26.50)	0.0020	10.60 (-11.65-32.85)	0.5007
Q2	23.00 (8.24-37.76)		25.00 (2.12-47.88)		18.91 (9.26-28.57)		15.36 (-5.92-36.64)	
Q3	20.00 (5.53-34.47)		28.00 (3.16-52.84)		16.50 (4.56-28.44)		15.24 (-2.25-32.73)	
Q4	11.00 (-2.29-24.29)		16.00 (-9.39-41.39)		7.86 (-3.99-19.70)		8.04 (-12.60-28.68)	
Stage unknown								
Intercept	137.00 (126.08-147.92)	<.0001	428.00 (413.84-442.16)	<.0001	104.02 (83.60-124.45)	<.0001	387.05 (356.17-417.93)	<.0001
Q1	8.00 (-9.88-25.88)	0.9239	7.00 (-14.73-28.73)	0.6671	6.85 (-11.91-25.61)	0.8381	10.92 (-17.67-39.51)	0.3793
Q2	1.00 (-13.60-15.60)		6.00 (-20.42-32.42)		-1.34 (-17.12-14.44)		18.92 (-8.30-46.13)	
Q3	2.39 (-14.14-18.91)		6.00 (-12.65-24.65)		2.13 (-16.65-20.92)		4.00 (-19.00-27.00)	
Q4	3.00 (-11.56-17.56)		-5.00 (-24.83-14.83)		5.19 (-12.13-22.51)		-7.00 (-35.49-21.49)	

Chapter 7 : General discussion

7.1 Summary of Findings

The work presented in this thesis focuses on income inequities in colorectal cancer survival and the diagnostic interval. Several conclusions can be made from the findings. First, in colorectal cancer studies, neighbourhood income is a very poor proxy for individual income, resulting in an underestimation of the effects of income on survival (Manuscript 1). Second, in the absence of individual income measures or when individual income is only available for a fraction of the population, cancer researchers can use probabilistic bias analysis to provide bias-adjusted effect estimates on survival (Manuscript 2). Finally, despite limitations in accessing individual-level income measures in administrative data, experiencing low income is associated with longer diagnostic intervals, up to 15 additional days, compared to individuals experiencing high income (Manuscript 3). Since neighbourhood income often underestimates individual income, income may have an even greater effect on the diagnostic interval at the individual level.

7.2 Implications and opportunities for future research

The collective results of my thesis drive home the conclusion that income inequities are ever-present in our cancer care system. Whether measured at the individual- or neighbourhood level, individuals experiencing low income consistently have worse cancer survival and longer diagnostic intervals. While I examined only these two outcomes for my thesis, these inequities extend to outcomes throughout the cancer care continuum.[1–3] The next steps in understanding and intervening on income inequities within the cancer system are to examine questions that explain pathways through which individuals with low income have worse cancer outcomes and

to evaluate programs and policies from an equity perspective that considers individuals experiencing vulnerabilities.

Tackling the issue of cancer care income inequities is complex and multifaceted. Environmental, structural, economic, historical, and cultural factors all work together to impact pathways through which individuals with low-income experience worse cancer outcomes. Interventions and policies that target the root cause of income inequities may be the most effective in improving equity in the cancer system.[4] For example, policies that target income directly, such as programs for guaranteed basic income or supplemental income, aim to address the unequal distribution of power and access to resources and can improve health outcomes such as birth weight and mental health.[5] Currently, two guaranteed income projects are being piloted in the US to determine if they can improve outcomes such as incidence, treatment adherence and quality of life in people with cancer.[6,7]

While programs that target inequities at the root may be the most effective at improving equity in the cancer system, this does not discount programs that aim to specifically improve income inequities within the cancer system. Evaluating current programs that are in place to improve access to cancer care from an equity lens would help in understanding how these programs help individuals experiencing vulnerabilities. For example, future research could compare outcomes for programs that aim to improve the diagnostic interval between income groups, such as the diagnostic assessment programs implemented in Ontario. These programs have been evaluated in the overall cancer populations, but it has yet to be determined if they work equally among individuals with low and high income.[8,9]

One of the main contributions of this thesis is describing and providing a solution for the misclassification of individual income by using adjusted neighbourhood income. Access to

individual income data remains difficult in Canada and other countries, such as the US and the UK. While Statistics Canada now provides access to linked datasets, including tax files, through the research data centres (RDC), there are many barriers to accessing this type of data, including administrative red tape and no remote access. Furthermore, the healthcare data at the RDC does not include physician billing data, which is imperative for defining outpatient visits and diagnostic tests. While I provide one solution to estimating the effects of individual income on survival when only neighbourhood income is available by using probabilistic bias analysis, this solution cannot be applied to cohorts outside of Canadian colorectal cancer patients and for outcomes other than survival. This limitation was played out in my third objective, where I had to use neighbourhood income instead of individual income due to the nature of the ICES administrative data. As we try to understand the effect of income on outcomes and how to intervene, Canada and other countries must start working towards solutions to provide data based on individual-level income to researchers using routinely collected linked data. In the interim, future research could build off my method described in objective two by determining if bias parameters from the whole country can be applied to sub-populations, such as those with cancer.

7.3 Strengths

Using linked routinely collected data for all three manuscripts in this thesis provided detailed population-based information on Canadian and Ontarian individuals with CRC. The data at ICES, in particular, allows us to capture a cohort of almost all individuals diagnosed with CRC during the study period. This enables us to make conclusions that apply to the population of CRC patients.

I used a conceptual model for all three manuscripts to determine appropriate confounders and avoid controlling for variables that might be on the causal pathway. Effects of main equity

exposures often occur jointly with other equity stratifiers, and controlling for these variables can decompose and diminish the main effect.[10,11] For example, if we examine the effect of income on cancer survival, we might consider controlling for comorbidities. However, people with low income disproportionally experience barriers to healthcare even before their cancer diagnosis, resulting in a higher prevalence of comorbid conditions, which can result in worse survival compared to individuals with high income. Controlling for comorbidities would decompose the effect of income on survival through comorbidities rather than giving an accurate picture of the unjust differences by income.[10,11] In this thesis, I aimed to describe true unjust income inequalities in survival and the diagnostic interval by avoiding controlling for causal pathway variables that would decompose the effect.

7.4 Limitations

Specific limitations are noted in each of the three manuscripts. Here, I consider the overarching limitations of the thesis as a whole.

While routinely collected data provided us with a representative cohort of Canadian individuals with CRC, it has some limitations. Most importantly, the availability of individual socioeconomic variables and comprehensive health administrative data in one place. While the RDC data does provide access to individual-level census- and tax-derived income variables, it does not have these variables linked to physician billing data. Moreover, while ICES has access to physician billing and hospitalization data, it does not have linked individual-level socioeconomic variables. This means that I could not measure the diagnostic interval in the RDC data to either calculate the diagnostic interval directly using the Statistics Canada data or use quantitative bias analysis to determine the effects of individual income on the diagnostic interval and apply that to the ICES data.

All three manuscripts might suffer from selection bias. Individuals who are most at risk for low income might not be captured in either the CanCHEC or the ICES data. The long-form census does not include people who are institutionalized at the time of the census, for example, people who reside in nursing homes, penitentiaries, and group homes. Furthermore, I excluded individuals with missing income information, i.e. those who either did not self-report their income or did not consent to the administrative linkage of their tax files. These people may be more likely to experience low income and possibly also more likely to experience poor cancer outcomes. In the ICES data, individuals who experience houselessness or never received a cancer diagnosis are not represented in our cohort. These people might be more likely to have low income and to experience poor cancer outcomes. Both these scenarios may lead to an underestimation of the number of people experiencing low income and having worse outcomes, potentially resulting in attenuation of the estimated effects.

I use individual and household income interchangeably throughout this thesis when what is being measured most often is household income. I chose to use household income over individual income for two reasons. First, it is the measure that Statistics Canada uses to calculate their neighbourhood income variable; therefore, household income is the closest measure to neighbourhood income and provides the best comparison. Second, household income considers the spending capabilities of the household in which people live, where individuals usually share resources. However, it is important to recognize that household income is not individual income and has its limitations. Most notably, if the wealth is not shared equally within the household, this may result in misclassification where income is over-represented for some individuals. This likely affects women, who receive a smaller share of their income from employment than men and may be at greater risk of having limited influence over important household economic

decisions.[12] Using household income instead of individual income might bias my results toward the null since I would be overestimating the income of certain household members. However, given the context of this study within Canada, I believe that household income more accurately describes an individual's spending power.

7.5 Conclusion

This thesis demonstrates income inequities in survival and the diagnostic interval among individuals with CRC in Canada. The results from this thesis also stress the importance of measuring individual income in order to accurately estimate income inequalities in cancer outcomes and provides a potential solution to estimate individual-level income effects when only neighbourhood income is available. As we move from understanding income inequities in cancer care to implementing and evaluating policies and programs to improve inequities and outcomes for marginalized populations, it is imperative that we have appropriate measurements of income. To accurately assess policies and programs to improve cancer outcomes for individuals experiencing inequities, data custodians should work towards providing researchers access to linked administrative data, including measures based on individual-level socioeconomic data.

7.6 References

1. Blair A, Gauvin L, Ouédraogo S, Datta GD. Area-Level Income Disparities in Colorectal Screening in Canada: Evidence to Inform Future Surveillance. Current Oncology. 2019 Apr;26(2):128–37.

2. Decker KM, Lambert P, Bravo J, Demers A, Singh H. Time Trends in Colorectal Cancer Incidence Rates by Income and Age at Diagnosis in Canada From 1992 to 2016. JAMA Network Open. 2021 Jul 19;4(7):e2117556.

3. Booth CM, Li G, Zhang-Salomons J, Mackillop WJ. The impact of socioeconomic status on stage of cancer at diagnosis and survival: a population-based study in Ontario, Canada. Cancer. 2010 Sep 1;116:4160–7.

4. Sayani A. Health Equity in National Cancer Control Plans: An Analysis of the Ontario Cancer Plan. Int J Health Policy Manag. 2019 Jun 16;8(9):550–6.

5. Gibson M, Hearty W, Craig P. The public health effects of interventions similar to basic income: a scoping review. The Lancet Public Health. 2020 Mar 1;5(3):e165–76.

6. Conger K. Stanford Medicine-led study on income supplementation, cancer inequities in persistent poverty areas. Stanford Medicine News Center [Internet]. 2022 May 26 [cited 2023 Jul 27]; Available from: http://med.stanford.edu/news/all-news/2023/06/cancer-poverty-grant.html

7. Doherty M, Heintz J, Leader A, Wittenberg D, Ben-Shalom Y, Jacoby J, et al. Guaranteed Income and Financial Treatment (G.I.F.T.): a 12-month, randomized controlled trial to compare the effectiveness of monthly unconditional cash transfers to treatment as usual in reducing financial toxicity in people with cancer who have low incomes. Front Psychol. 2023 May 18;14:1179320.

8. Habbous S, Khan Y, Langer DL, Kaan M, Green B, Forster K, et al. The effect of diagnostic assessment programs on the diagnosis and treatment of patients with lung cancer in Ontario, Canada. Ann Thorac Med. 2021;16(1):81–101.

9. Habbous S, Khan Y, Green B, Martin T, Kaan M, Kennedy E, et al. The impact of diagnostic assessment programs on the diagnosis and treatment of colon and rectal cancers in a single-payer system. American Journal of Preventive Medicine and Public Health. 2020 Jun 20;6(4):80–108.

10. Jackson JW. Meaningful causal decompositions in health equity research: definition, identification, and estimation through a weighting framework. Epidemiology. 2021 Mar 1;32(2):282–90.

11. VanderWeele TJ, Robinson WR. On causal interpretation of race in regressions adjusting for confounding and mediating variables. Epidemiology. 2014 Jul;25(4):473–84.

12. De Henau J, Himmelweit S. Unpacking Within-Household Gender Differences in Partners' Subjective Benefits From Household Income. Journal of Marriage and Family. 2013;75(3):611–24.