Social contexts of survival: unraveling lung cancer inequalities in Canada

Samia Qureshi, MScPH

Department of Epidemiology and Biostatistics
McGill University, Montréal

December 30, 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Epidemiology

Table of Contents

List of tables

List of figures

List of acronyms and abbreviations

area under the receiver operating characteristic curve (AUC)

Canada Child Tax Benefit (CCTB)

Canada Revenue Agency (CRA)

Canadian Cancer Registry (CCR)

Canadian Cancer Registry tabulation master file (CCR TMF)

Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures (CCP)

Canadian Classification of Health Interventions (CCI)

Canadian Community Health Survey (CCHS)

Canadian Institutes of Health Research (CIHR)

Canadian Partnership Against Cancer (CPAC)

Canadian Research Data Centre Network (CRDCN)

Canadian Vital Statistics Database (CVSD)

Census Agglomerations (CA)

Census Metropolitan Areas (CMA)

confidence interval (CI)

coronavirus disease 2019 (COVID-19)

Derived Record Repository (DRD)

directed acyclic graph (DAG)

Discharge Abstract Database (DAD)

dissemination area (DA)

drug identification numbers (DIN)

epidermal growth factor receptor (EGFR)

epidermal growth factor receptor tyrosine kinase inhibitor (EGFR-TKI)

Fonds de recherche du Québec – Santé (FRQS)

Fichier d'inscription des personnes assurées (FIPA)

Institut national d'excellence en santé et en services sociaux (INESSS)

Institut universitaire de cardiologie et de pneumologie de Québec – Université Laval (IUCPQ-UL)

International Classification of Diseases, 10th Revision, with Canadian Enhancements (ICD-10-CA)

International Classification of Diseases for Oncology, third edition (ICD-O-3)

international nonproprietary names (INN)

interquartile range (IQR)

intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy (I-MAIHDA)

inverse-probability-of-treatment weight (IPTW)

Longitudinal Immigration Database (IMDB)

Maintenance et exploitation des données pour l'étude de la clinetèle hospitalière (MED-ECHO)

median odds ratio (MOR)

Metropolitan Influence Zone (MIZ)

Ministère de la Santé et des Services sociaux (MSSS)

National Ambulatory Care Reporting System (NACRS)

National Prescription Drug Utilization Information System (NPDUIS)

Organization for Economic Cooperation and Development (OECD)

Postal Code Conversion File – plus (PCCF+)

proportional change in variance (PCV)

pure natural direct effect (PNDE)

Quebec Inter-University Centre for Social Statistics (QICSS)

Régie de l'assurance maladie du Québec (RAMQ)

Research Data Centre (RDC)

Services pharmaceutiques (SMED)

Services rénumérés à l'acte (SMOD)

Social Data Linkage Environment (SDLE)

socioeconomic status (SES)

sociodemographic status (SDS)

Statistical Area Classification (SAC)

Surveillance, Epidemiology, and End Results (SEER)

Sustainable Development Goals (SDG)

T1 Family File (T1FF)

T1 Personal Master File – Death (T1PMF-D)

total effect (TE)

variance partition coefficient (VPC)

natural indirect effect (TNIE)

World Health Organization (WHO)

Abstract

Lung cancer is one of the leading causes of cancer deaths in Canada. While income-based disparities in lung cancer survival are well documented, the influence of other social determinants—such as education, immigration status, and place of residence—remains poorly understood. This knowledge gap stems from limited data on social factors in cancer registries, as highlighted by the Canadian Cancer Statistics Advisory Committee's 2024 report. My thesis aims to address this gap by analyzing linked databases that combine socioeconomic, sociodemographic, and geographic information with cancer data. It aligns with global efforts like Sustainable Development Goal 3, which seeks to reduce premature mortality from non-communicable diseases, including cancer, by one-third by 2030. My research also supports the Canadian Partnership Against Cancer's strategy to promote equity in the cancer system by 2029.

In the first manuscript of my thesis, I examined social disparities using data from the nationally representative Canadian Community Health Survey linked to the Canadian Cancer Registry and income tax records. This linkage allowed me to compare lung cancer survival across individual-level measures of family income, education, immigration status, and the urban-rural spectrum of place of residence. I applied Kaplan-Meier analysis and additive hazards models to estimate survival inequalities among 4,430 survey respondents diagnosed with primary lung cancer between 2010 and 2016. Survival differences emerged across the urban-rural spectrum (Kaplan-Meier logrank p-value = 0.0215). We observed 50.1 additional deaths per 10,000 person-months (95% CI: -3.7 to 107.2) in small urban cities relative to large urban cities, albeit with some imprecision. Point estimates from mediation analysis showed that 41% of this effect could be attributed to advanced stage at diagnosis, which did not appear to contribute to survival inequalities based on socioeconomic status.

Next, I used Quebec's health administrative data to study inequalities in survival and treatment among advanced lung cancer patients based on area-level income, education, immigration, material deprivation, and geographic location. The study included 457 patients who received gefitinib, an EGFR tyrosine kinase inhibitor (EGFR-TKI) known as a targeted therapy, as first-line palliative treatment between 2001 and 2019. Although imprecise, our estimate indicates that the median survival time among patients residing in areas with the highest levels of material deprivation was 0.69 (95% CI: 0.47–1.04) times that of those in the least deprived areas. Compared to individuals living in Montreal, the probability of receiving osimertinib, another EGFR-TKI, as a second-line palliative treatment was lower for those in other large urban areas (relative ratio: 0.39, 95% CI: 0.16–0.71) and in rural regions

(relative ratio: 0.55 (95% CI: 0.25–0.98). Additionally, individuals residing in health regions peripheral to Quebec or Montreal experienced longer median wait times for using gefitinib as first-line palliative treatment - 45 days compared to 35 in regions with university-affiliated health centres (relative ratio of 1.27, 95% CI: 1.09–1.54).

For my third study, I used the Canadian Cancer Registry linked to income tax and immigration data to estimate population-level survival for lung cancer patients at the intersection of four social identities (age, sex, income, and immigration status). I used multilevel modeling to predict 1-year survival probabilities for 183,350 individuals diagnosed between 2002 and 2013. I observed substantial variation in 1-year survival probabilities across 72 intersectional strata, ranging from 23.3% to 66.9%. Among individuals aged <60, low-income men who were long-term residents had the lowest 1-year survival probability at 36.3% (95% CI: 34.44 to 38.19), which was comparable to that of low-income long-term resident men aged 70–74 (33.9%, 95% CI: 32.06 to 35.55). Their survival probability was also 2.8 percentage points (95% CI: -5.0 to -0.5) lower than expected based solely on the additive effects of age, sex, income, and immigration status. Low-income immigrant women aged ≥80, who could have been in Canada the longest compared to other immigrant women in the study, were also found to experience 2.8 percentage points (95% CI: -5.4 to -0.3) lower than expected 1-year survival based on additive effects alone. Most intersectional effects diminished after accounting for stage at diagnosis.

Overall, my thesis provides a comprehensive understanding of lung cancer survival inequalities in Canada by incorporating detailed measures of social factors. I demonstrate that survival varies by socioeconomic status, sociodemographic status, and urbanicity, with disparities also evident at advanced stages of lung cancer in the era of breakthrough therapies. The thesis further underscores differences in stage at diagnosis and access to treatment based on place of residence. Importantly, it emphasizes the need to consider intersecting social identities and identifies key populations—such as young, low-income, long-term resident men and immigrant women, particularly those who have been in Canada for several decades—with unique survival disadvantages. Addressing challenges in early diagnosis and access to innovative treatments for individuals with social disadvantages in lung cancer survival could help reduce lung cancer mortality in Canada.

Resumé

Le cancer du poumon est l'une des principales causes de décès liés au cancer au Canada. Bien que les disparités dans la survie de cancer du poumon basées sur le revenu soient bien documentées, l'influence d'autres déterminants sociaux—tels que l'éducation, le statut d'immigration et le lieu de résidence—reste méconnue. Cette lacune de connaissances découle de la disponibilité limitée de données sur les facteurs sociaux dans les registres de cancer, comme l'a souligné le rapport 2024 du Comité consultatif des statistiques canadiennes sur le cancer. Ma thèse vise à combler cette lacune en analysant des bases de données liées qui combinent des informations socioéconomiques, sociodémographiques, et géographiques avec des données sur le cancer. Elle s'inscrit dans les efforts mondiaux tels que l'objectif de développement durable 3, qui cherche à réduire d'un tiers la mortalité prématurée due aux maladies non transmissibles. Ma recherche soutient également la stratégie du Partenariat canadien contre le cancer visant à promouvoir l'équité dans le système de lutte contre le cancer d'ici 2029.

Dans le premier manuscrit de ma thèse, j'ai examiné les disparités en utilisant les données de l'Enquête sur la santé dans les collectivités canadiennes, qui est représentative à l'échelle nationale, liées au Registre canadien du cancer et aux dossiers fiscaux. Ce couplage de données m'a permis de comparer la survie de cancer du poumon selon des mesures individuelles de revenu familial, niveau d'éducation, et de statut d'immigration, ainsi qu'un spectre urbain-rural du lieu de résidence. J'ai appliqué des analyses de Kaplan-Meier et de modélisation multiniveaux pour estimer les inégalités de survie chez 4,430 répondants de l'enquête ayant reçu un diagnostic primaire de cancer du poumon entre 2010 et 2016. Des différences de survie ont été observées selon le spectre urbain-rural (p-valeur logrank de Kaplan-Meier = 0,0215). L'analyse de médiation a également montré que 41 % de la différence de risque entre les individus résidant dans de petites villes urbaines et ceux des grandes villes urbaines (50,1 décès supplémentaires pour 10 000 personnes-mois, IC 95 % : -3,7 à 107,2) pouvait être attribué à un stade avancé au diagnostic, qui ne semblait pas contribuer aux inégalités de survie selon le statut socio-économique.

Ensuite, j'ai utilisé les données clinico-administratives du Québec pour étudier les inégalités de survie et de traitement chez les patients atteints de cancer du poumon avancé, en fonction du revenu, du niveau d'éducation, du statut d'immigration et de défavorisation matérielle au niveau

du quartier et de situation géographique. L'étude a inclus 457 patients ayant reçu le géfitinib, un inhibiteur de la tyrosine kinase de l'EGFR (EGFR-TKI) connue comme thérapie ciblée, en première ligne de traitement palliatif entre 2001 et 2019. Les patients résidant dans les zones les plus touchées par la défavorisation matérielle avaient une survie médiane de 15,1 mois, soit 0,69 (IC 95 % : 0,47–1,04) fois celle des patients des zones les moins défavorisées. Comparativement aux individus résidant à Montréal, la probabilité de recevoir de l'osimertinib, un autre EGFR-TKI, en deuxième ligne de traitement palliatif était plus faible pour ceux vivant dans d'autres grandes régions urbaines (ratio relatif : 0,69, IC 95 % : 0,47–1,04) et dans les régions rurales (ratio relatif : 0,55, IC 95 % : 0,25–0,98). De plus, les individus résidant dans des régions socisanitaires périphériques à Québec ou Montréal ont connu des délais d'attente médians plus longs pour l'utilisation du géfitinib en première ligne de traitement palliatif, avec un retard de 45 jours contre 35 dans les régions avec des centres hospitaliers universitaires (ratio relatif : 1,27, IC 95 % : 1,09–1,54).

Pour ma troisième étude, j'ai utilisé le Registre canadien du cancer, lié aux données fiscales et d'immigration, pour estimer la survie à l'échelle de la population des patients atteints de cancer du poumon à l'intersection de quatre identités sociales (âge, sexe, revenu et statut d'immigration). J'ai utilisé la modélisation multiniveau pour prédire les probabilités de survie à 1 an pour 183,350 individus diagnostiqués entre 2002 et 2013. J'ai observé une variation substantielle a été retrouvée à travers les 72 strates intersectionnelles dans les probabilités de survie à 1 an, allant de 23,3 % à 66,9 %. Parmi les individus âgés de moins de 60 ans, les hommes à faible revenu qui étaient résidents de Canada depuis longtemps avaient une probabilité de survie à 1 an la plus faible, à 36,3 % (IC 95 % : 34,44, 38,19), comparable à celle des hommes à faible revenu qui étaient aussi résidents depuis longtemps mais âgés de 70 à 74 ans (33,9 %, IC 95 % : 32,06, 35,55). Leur probabilité de survie était également inférieure de 2,8 points de pourcentage (IC 95 % : -5,0 à -0,5) à celle attendue, basée uniquement sur les effets additifs de l'âge, du sexe, du revenu et du statut d'immigration. Les femmes immigrantes à faible revenu âgées de ≥80 ans, qui pouvaient être au Canada le plus longtemps possible par rapport aux autres femmes immigrantes dans l'étude, ont également présenté une probabilité de survie à 1 an inférieure de 2,8 points de pourcentage (IC 95 % : -5,4 à -0,3) par rapport à celle basée uniquement sur les effets additifs. La plupart des effets intersectionnels se sont estompés après avoir pris en compte le stade au moment du diagnostic.

En somme, ma thèse offre une compréhension approfondie des inégalités de survie de cancer du poumon au Canada en intégrant des mesures détaillées des facteurs sociaux. Je démontre que la survie varie selon le statut socioéconomique, le statut sociodémographique et l'urbanisation, avec des disparités également évidentes aux stades avancés du cancer du poumon dans l'ère des thérapies innovantes. La thèse met en lumière des différences dans le stade au moment du diagnostic et l'accès au traitement en fonction du lieu de résidence. Elle souligne l'importance de prendre en compte les identités sociales croisées et identifie des populations clés — telles que les jeunes hommes à faible revenu qui sont résidents du Canada depuis longtemps, et les femmes immigrantes, en particulier celles qui sont au Canada depuis plusieurs décennies — confrontées à des désavantages uniques en termes de survie. Relever les défis du diagnostic précoce et de l'accès aux traitements innovants pour les personnes socialement défavorisées en matière de survie au cancer du poumon pourrait contribuer à réduire la mortalité liée au cancer du poumon au Canada.

Acknowledgements

First and foremost, I thank God for guiding me to this moment and helping me overcome countless obstacles. Your blessings throughout this PhD journey have been beyond what I could have imagined. I dedicate this thesis to my late father, who would have been so proud. Not a day goes by that I do not think of you—I miss you dearly.

To my supervisor, Dr. Erin Strumpf, thank you for your steadfast support throughout my journey at McGill. You have witnessed my failures, successes, moments of joy, and countless buckets of tears. Through it all, you stood by my side with patience, grace, and wisdom. Thank you for guiding me, allowing me to explore, pushing me a little when necessary, and gently pulling me back to realistic goals. I take pride in how far we've come, and I will always cherish you as a supervisor, colleague, and friend.

I want to thank the cancer team at INESSS for welcoming me as a Health System Impact fellow and providing me with experiences that set me apart as a student. I extend special thanks to Dr. Meghan McMahon, whose exceptional guidance and support as fellowship director were instrumental to my growth. I also thank Dr. Arijit Nandi and Dr. Mireille Snitzer for their invaluable feedback for advancing my manuscripts.

I am deeply grateful to the administrative staff at McGill's Epidemiology and Biostatistics Department—Katherine Hayden and André Yves Gagnon—and Danielle Forest at McGill's Laboratory of the Quebec Inter-University Centre for Social Statistics for always addressing my last-minute questions and requests. I could not have made it without their help and kindness.

To my husband and kids, you are my light, grounding force, and home. Your unconditional love has carried me through both difficult and joyful moments. Nabil, I am forever indebted to you for all the times you took care of everything and everyone while I spent my evenings at the desk. Thank you for being my rock; your belief in me has been my greatest strength. Yaqoob and Zahra, you have constantly motivated me to show that it's never too late to pursue your dreams and that hard work truly bears fruit. I look forward to all that life has in store for us next.

Mama, your strength and resilience despite life's challenges have been my source of inspiration. You have always entertained my wild ideas and been the first to push me to new heights. I love

you so much. You, Asma, Assad, and Usmaan have come through countless times, rain or shine, be it lifts, babysitting, food, or more. You are my roots, my village. What is life without you?

Finally, a shoutout to my squad: Sheerose, Razia, Tarannum, Carla, Doris, Arezoo, and Mariam! You've been my confidantes, cheerleaders, and partners in crime. Thank you for always being there to listen, laugh, and lift me up. I'm lucky to have you as friends.

Statement of financial support

I am deeply grateful for the financial support provided by McGill University, the Canadian Institutes of Health Research (CIHR - Institutes of Health Services and Policy Research, Population and Public Health, and Cancer Research), the Institut national d'excellence en santé et en services sociaux (INESSS), the Fonds de recherche du Québec – Santé (FRQS), the Quebec Inter-University Centre for Social Statistics (QICSS), and the Canadian Research Data Centre Network (CRDCN).

I received two CIHR awards: the Health System Impact Fellowship (co-funded by INESSS, 2019–2020) and the Frederick Banting and Charles Best Canada Graduate Scholarship (2021–2024). I also received a provincial doctoral award from the FRQS (2021-2024: declined May 2021 and onwards), a QICSS fellowship and conference grants (2024), and the CRDCN Emerging Scholars Grant (2024). These awards supported my dissertation, which used Statistics Canada data accessed through the QICSS McGill-Concordia lab.

I am particularly thankful to my supervisor, Dr. Erin Strumpf, for the supplemental funding provided throughout the academic years, including 2024-2025, enabling me to continue my research and publish my second manuscript in an open-access journal.

This generous support allowed me to focus on my research without financial stress, significantly contributing to the successful completion of my doctoral degree.

Statement of originality

The three manuscripts in this thesis were guided by the research questions I formulated and the study methods I proposed. I conducted the studies described in each manuscript, while incorporating feedback from my supervisor, committee member, and co-authors at various stages of development, as acknowledged in each manuscript.

The goal of my thesis is to fill in some of these gaps to advance work on equity in lung cancer survival and care in Canada. The ideas for my thesis were informed by gaps identified in the literature regarding social and geographic inequalities in lung cancer survival in Canada, including the lack of national-level estimates, limited use of individual-level or more nuanced measures of social position, the absence of formal assessments of stage at diagnosis as a mediator, the impact of next-generation therapies for advanced cancers, and the role of intersecting social identities. In Manuscript 1, I incorporated individual-level measures of socioeconomic and sociodemographic status, as well as more detailed urban-rural classifications, to provide national-level estimates of inequalities in lung cancer survival and to quantify the role of stage at diagnosis. In Manuscript 2, I estimated social and geographic inequalities in lung cancer survival within the context of next-generation treatments (e.g., targeted therapies) for advanced lung cancer, and measured access to these treatments (use and delays). The findings from these two manuscripts further inspired me to quantify national-level inequalities in lung cancer survival across intersecting social identities measured at the individual level, which formed the basis of Manuscript 3.

Below are two INESSS reports and three peer-reviewed publications I authored or co-authored during my doctoral training that are within the scope of cancer research but beyond the work presented in this thesis.

Published

- **Qureshi, S.,** Boily, G., Boulanger, J., Golo, K. T., Guédon, A. -C., Lehuédé, C., Roussafi, F., Truchon, C., & Strumpf, E. (2022). Advanced Lung Cancer Patients' Use of EGFR Tyrosine Kinase Inhibitors and Overall Survival: Real-World Evidence from Quebec, Canada. *Current Oncology*, *29*(11), 8043-8073. https://doi.org/10.3390/curroncol29110636

- Institut national d'excellence en santé et en services sociaux (INESSS). Utilisation en contexte québécois des inhibiteurs de la tyrosine kinase du récepteur du facteur de croissance épidermique (EGFR) pour le traitement du cancer du poumon. État des pratiques rédigé par Gino Boily, Kossi Thomas Golo, Aude- Christine Guédon, Camille Lehuédé, **Samia Qureshi**, Ferdaous Roussafi et Erin Strumpf. Québec, Qc2022. p. 109.

- Boily G, Guédon A-C, Golo KT, **Qureshi S**, Lehuédé C, Strumpf E. Création et caractérisation d'une cohorte québécoise de patients atteints d'un cancer du poumon à l'aide de données clinico-administratives Québec (Québec): Institut national d'excellence en santé et en services sociaux (INESSS); 2021.

- Datta, G. D., Mayrand, M. H., **Qureshi, S**., Ferre, N., & Gauvin, L. (2020). HPV Sampling Options for Cervical Cancer Screening: Preferences of Urban-Dwelling Canadians in a Changing Paradigm. *Current Oncology*, *27*(2), 171-181. https://doi.org/10.3747/co.27.5089

- L'Espérance, K., Datta, G. D., **Qureshi, S**., & Koushik, A. (2020). Vitamin D Exposure and Ovarian Cancer Risk and Prognosis. *International Journal of Environmental Research and Public Health*, *17*(4), 1168. https://doi.org/10.3390/ijerph17041168

Contribution of authors

For each of the three studies listed below, I conceptualized the study, curated the data, conducted the analyses, compiled the results, and prepared the manuscript, incorporating revisions from co-authors. My supervisor, Dr. Erin Strumpf, played a key role in guiding me during the planning of these studies and providing recommendations that I applied at all stages, including my dissertation.

Manuscript 1: Qureshi S, Schnitzer M, Nandi A, Strumpf E. Social Inequalities in Survival Among Canadian Lung Cancer Patients and the Mediating Role of Stage at Diagnosis. *To be submitted.*

Dr. Erin Strumpf, Dr. Mireille Schnitzer, and Dr. Arijit Nandi guided the interpretation of the results and provided revisions to the manuscript I drafted.

Manuscript 2: Qureshi S, Boily G, Boulanger J, Pagé É, Strumpf E. Inequalities in survival and care across social determinants of health in a cohort of advanced lung cancer patients in Quebec (Canada): A high-resolution population-level analysis. *Cancer Med*. 2023;12(11):12683-704. DOI: 10.1002/cam4.5897

The study presented in Manuscript 2 was conducted in collaboration with the cancer team INESSS, during a Health System Impact Fellowship that I completed at that institution. My supervisor, Dr. Erin Strumpf, facilitated this study, and Élisabeth Pagé provided the necessary resources, including desk space and secure access to data. Dr. Erin Strumpf, Dr. Gino Boily, and Dr. Jim Boulanger guided the development of the methods and validated the results. All co-authors provided revisions and approved the revised manuscript.

Manuscript 3: Qureshi S, Nandi A, Strumpf E. Age, Gender, Income and Immigration Status and Lung Cancer Survival in Canada: A Population-Level Intersectional Multilevel Analysis. *To be submitted.*

Dr. Erin Strumpf and Dr. Arijit Nandi provided guidance on interpreting the results and revisions for the manuscript I drafted.

Chapter 1. Introduction

1.1. Background

Among the 17 United Nations Sustainable Development Goals (SDGs) adopted in 2015 to transform lives globally by 2030, Goal 3 is to ensure healthy lives and promote well-being for all people of all ages.[1] Although this goal may appear less urgent for developed countries, which generally perform better than the global average, specific populations within these nations continue to face barriers to healthcare and poorer health outcomes.

Health equality involves achieving some equal measure of full, healthy lives for all individuals.[2] However, since individual needs vary, health equity—where everyone is at their fullest health potential—can only be achieved by proportional allocation of resources based on these needs. Inequities arise when systemic, avoidable, and unfair circumstances prevent some individuals from reaching this potential.[2] Identifying inequalities is a fundamental step in taking action to advance health equity and ensuring that everyone has the opportunity to achieve optimal health outcomes. In 2018, the first pan-Canadian effort was made to document inequalities in health outcomes in Canada.[3] This initiative focused on 22 indicators of health status and social, economic, and environmental conditions that impact health. The analyses revealed inequalities among key populations in Canada, who are defined by socioeconomic status (SES), sociodemographic status (SDS), and place of residence, despite the country's universal healthcare system providing free access at the point of care. Furthermore, the report acknowledged that health inequalities are driven by the complex and compounding influence of multiple social identities, highlighting the need to explore how they intersect to shape health outcomes.

SDG targets under Goal 3 focus on reducing premature mortality due to non-communicable diseases by one-third through different ways, including prevention and treatment. Cancer is the second leading cause of premature deaths globally, following cardiovascular diseases.[4, 5] In high-income countries, it is the leading cause of premature deaths, with lung cancer being the most significant contributor.[6, 7] The 2018 pan-Canadian analysis on inequalities did report inequalities in lung cancer incidence; however, no measures related to death (i.e., premature mortality or survival) were provided for a more comprehensive overview of the cancer burden.[3] Canada's 2021 annual report on the 2030 Agenda and SDGs does note an overall improvement in

reducing premature mortality from all cancers.[8] However, the report does not include detailed data or disaggregated analysis to show how this progress varies across different populations, such as those defined by socioeconomic or demographic factors. The SDG indicator 3.4.1, established by the Inter-Agency and Expert Group on SDG Indicators to monitor progress toward SDG target 3.4, is specified as the "mortality rate attributed to cardiovascular disease, cancer, diabetes, and chronic respiratory disease" with data disaggregated by "income, sex, age, race, ethnicity, migratory status, disability, and geographic location" when applicable.[9]

The 10-year cancer control strategy set by the Canadian Partnership Against Cancer (CPAC) aims to enhance the cancer care system and improve cancer outcomes in Canada by 2029.[10] Among the five strategic priorities outlined in the plan, one key area is eliminating barriers to care for underserved communities. Another priority is ensuring that cancers are diagnosed at earlier stages, which includes implementing nationwide lung cancer screening programs. A third priority focuses on delivering high-quality care by establishing and adopting best practices and standards, such as the use of genetic testing, which is essential for prescribing new treatments that target specific genetic mutations in cancer cells (i.e., targeted treatments). The introduction of advanced therapies like targeted treatments has significantly improved the prognosis for patients with advanced lung cancer. However, these therapies are costly, and access to them remains restricted, partly due to the absence widespread genetic testing and of a national universal public drug plan. The CPAC priorities align with global efforts, including SDGs targets.[1] Target 3.8, in particular, emphasizes achieving universal health coverage, encompassing "financial risk protection, access to quality essential health-care services, and access to safe, effective, quality, and affordable essential medicines and vaccines for all."[1] Furthermore, CPAC's strategy aligns with the 2022-2024 Close the Care Gap campaign launched on World Cancer Day by the Union for International Cancer Control, which aims to address the most pressing disparities in cancer care and control.[11, 12]

To ensure equitable actions under CPAC's strategic plan, it is essential first to identify individuals experiencing worse cancer outcomes and access to care along the cancer continuum relative to others. This thesis aims to investigate lung cancer survival inequalities in Canada across several unexplored measures of socioeconomic and sociodemographic status and place of residence, with a special focus on diagnosis and treatment. By understanding these patterns, this research seeks to

advance literature on lung cancer survival inequalities in Canada and, in return, inform future research, and policies and practices that align with CPAC's strategic priorities and global frameworks like the SDGs.

1.2. Research objectives

This manuscript-based thesis consists of three objectives, as follows:

> **Objective 1:** Assess inequalities in lung cancer survival across indicators of SES, SDS, and place of residence, while focusing on understanding the role of stage at diagnosis in explaining these disparities. (Manuscript 1)
>
> **Objective 2:** Investigate inequalities in lung cancer survival among patients with advanced lung cancer who are using targeted therapies. This includes assessing disparities in survival across SES, SDS, and place of residence and evaluating inequalities in access to these therapies. (Manuscript 2)
>
> **Objective 3:** Examine inequalities in lung cancer survival at the intersection of multiple dimensions of SES and SDS while considering the context of time and place of residence. (Manuscript 3)

1.3. Thesis structure

This thesis includes seven chapters. In this chapter (**Chapter 1**), I have introduced the topic of my thesis and my research objectives. In **Chapter 2,** I provide a comprehensive review of lung cancer epidemiology and inequalities in survival, diagnosis, and treatment. I also discuss theoretical frameworks and identify knowledge gaps that justify the three objectives. **Chapter 3** covers the data sources used for all three objectives. The three manuscripts addressing each objective are presented in the following three chapters**. In Chapter 4,** I examine inequalities in overall survival by income, education, immigration status, and urban-rural residence among individuals diagnosed with lung cancer who participated in a national health survey. Mediation analysis is applied to quantify the role of diagnosis at an advanced stage in these disparities. In **Chapter 5,** I leverage Quebec's health administrative data to investigate survival inequalities among advanced-stage lung cancer patients prescribed targeted therapies, focusing on area-level SES, immigration status,

and residence defined by census or health region boundaries. Inequalities in time to first-line targeted therapies and access to second-line treatments are also assessed. In **Chapter 6**, I apply quantitative intersectional multilevel analysis to assess survival across individuals with intersecting identities based on age, sex, income, education, and immigration status in the Canadian lung cancer population. Finally, in **Chapter 7,** I synthesize the findings, discuss strengths and limitations, and propose future directions for addressing inequities in lung cancer survival.

Chapter 2. Literature review

In this chapter, I provide an overview of lung cancer epidemiology, including etiology, disease burden (incidence, mortality, survival), tumour histology, and cancer staging. I then dive into inequalities in lung cancer survival and the social determinants of health explored in my manuscripts. I also discuss the theoretical frameworks related to these determinants that underpin my research.

2.1. Lung cancer epidemiology

2.1.1. Etiology

The Surgeon General's report in 1964 in the United States was one of the first to establish that smoking is one of the main risk factors for lung cancer.[13] Since the report, smoking rates have been declining in the United States and worldwide.[13] Although not addressed in the report, the effect of second-hand exposure to tobacco smoking has also been causally linked to lung cancer.[14] Other risk factors for lung cancer include genetic predisposition, diet and alcohol, chronic inflammation from infections, ionizing radiation like radon, occupational exposures like asbestos, and air pollution.[14] It is estimated that 74% of all lung cancer cases in 2020 were due to active tobacco smoking, and 85% were due to active/passive tobacco smoking, physical inactivity, residential radon, air pollution and poor diet altogether.[15]

2.1.2. Incidence and mortality

In Canada, lung cancer is the most commonly diagnosed cancer in men and women combined, and it is the second most frequently diagnosed cancer among each sex.[15] There are an estimated 15,300 and 15,800 incident cases of lung cancer in Canada in 2023, with age-standardized incidence rates of 60.1 cases per 100,000 men and 58.4 cases per 100,000 women.[15] Lung cancer accounts for 23% and 24% of all cancer deaths in men and women, respectively, making it the leading cause of cancer deaths in each sex.[15] There were 10,800 and 9,800 deaths due to lung cancer expected in 2023, with age-standardized mortality rates of 60.1 cases per 100,000 people and 58.4 cases per 100,000 people in men and women, respectively.[15]

Lung cancer is a disease that occurs at older ages, with incidence and mortality rates peaking at 75 years or more.[15] In general, the rates of lung cancer incidence and mortality in men and women have been following smoking patterns. For example, the incidence and mortality rates are slightly higher in men because they have always had higher smoking rates than women.[16] Furthermore, a decline in lung cancer incidence and mortality began about 20 years after the time when the prevalence of daily tobacco smoking started to decline, which was the mid-1960s for men and the mid-1980s for women.[15, 17] Although incidence and mortality rates have been decreasing with time, the number of incident cases and deaths due to lung cancer is expected to rise over time due to a growing and aging Canadian population.[15]

2.1.3. Histology and stage at diagnosis

Cancer of the lung and bronchus (lung cancer) is divided into different types based on the characterization of cancer cells under the microscope (i.e. histology). Non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) are the two main histologic types of lung cancers. NSCLC, which accounts for 88% of these histologically-defined cancers, is divided into three subtypes: adenocarcinoma, squamous cell carcinoma, and large cell carcinoma.[15] Adenocarcinoma is the most common histologic type of lung cancer, with better survival than other types, and its incidence has been increasing over time.[15, 18] In contrast, the incidence of other histologic types has been stagnant or decreasing over time.[15] The histologic types most strongly associated with smoking are squamous cell carcinoma and small cell carcinoma, which have the worst survival rates among all histologic types.[15, 18]

Cancer stage at diagnosis represents the extent of the disease based on tumour size and its spread in the body. It is assigned according to the tumour-node-metastases (TNM) rules defined by the American Joint Committee on Cancer (AJCC).[19] T is for describing the primary tumour, N is for describing the spread of the cancer to lymph nodes in surrounding regions of the primary tumour, and M is for describing the additional spread of the cancer in the rest of the body. There are four main types of lung cancer stages (I, II, III, and IV), and they are further divided into two subcategories (a and b) to reflect cancer prognosis, which worsens as stage increases.[20] In Canada, between 2012 and 2016, 21% of lung cancer cases were diagnosed at stage I, 8% at stage II, 20% at stage III, and 49% at stage IV.[21] The absence of a lung cancer screening program

during this period contributed to the high proportion of diagnoses at a later stage, highlighting the challenges of detecting patients with lung cancers at earlier stages in Canada. Ontario was the first province to launch a lung cancer screening program as a pilot initiative in 2017.[22] British Columbia is the first province to implement a province-wide organized lung cancer screening program in 2022.[23] Other provinces have also implemented pilot programs around the same time, and many are planning to roll out province-wide organized lung cancer screening programs.

The stages of lung cancer are grouped based on the potential for treatment with complete surgical removal of the tumour, which is considered a curative treatment.[24] Curative surgery alone is most effective for stage I-II NSCLC patients and is regarded as their standard of care. In contrast, stage IIIA NSCLC patients may benefit from combined surgery and chemoradiation.[25-27] Therefore, stage I, II, or IIIa NSCLC patients are considered "early-stage" as they may be amenable to curative surgery. In contrast, stage IIIb and IV NSCLC patients are considered "advanced-stage" and inoperable, with surgery offering no survival benefits.[25-27] Although surgery is possible for stage IIIa NSCLC patients, very few are resected in Ontario, with resection rates falling from 52.2 to 63.0% for stage I and II, respectively, to 26.4% for stage IIIa.[28] Therefore, patients with stage IIIa NSCLC are mostly incurable due to the extent of their disease, and stage IIIa patients are often grouped with stage IIIb and IV NSCLC patients to form the "advanced stage" category (i.e., locally advanced or metastatic). For SCLC patients, curative surgery may be beneficial for some patients with stage I to IIa disease; however, surgery is rarely seen in practice in Canada.[25, 29, 30] Although stage I to III SCLC patients are often grouped as "limited" disease, it has been recommended to separate stage III SCLC patients from this group due to significant differences in prognosis.[31]

2.1.4. Survival and treatment

Survival statistics allow us to understand how long individuals can survive once they are diagnosed with lung cancer. They are more likely to reflect inequities in cancer care relative to mortality statistics which may reflect inequities cancer risk (incidence) and/or cancer care. In Canada, survival from lung cancer remains the poorest among the top four most diagnosed cancers: lung, breast, prostate, and colorectal.[15] The 5-year survival probability for individuals diagnosed with lung cancer between 2015 and 2017 was 22%, whereas it was 91%, 89%, and 67% for those

diagnosed with prostate, female breast, and colorectal cancers, respectively.[15] This poor survival is most likely due to the fact that about 70% of all lung cancer patients are diagnosed at an advanced stage when 3-year survival probabilities range between 5 and 22%.[21] In general, the 5-year survival probability for individuals diagnosed with stage I lung cancer is 62%, in contrast to probabilities above 90% for patients diagnosed with stage I breast, prostate or colorectal cancer.[21]

The drastic drop in 3-year survival probabilities for lung cancer—from 71% for stage I to 5% for stages IV—is primarily linked to treatment options. At early stages, lung cancer can often be cured with surgery or surgery combined with chemoradiation therapy. However, when lung cancer is diagnosed at late stages, it cannot be cured, and chemotherapy is usually prescribed as the standard of care for controlling or managing the disease. Although chemotherapy relative to best supportive care extends the survival of lung cancer patients diagnosed at later stages, its effects are minimal.[32] Recent advancements in precision medicine have widened the scope of treatment options and contributed to improving advanced lung cancer patients' treatment outcomes and quality of life.[33] For example, targeted therapies are now used as a first-line palliative (non-curative) treatment for patients whose tumour cells harbor specific mutations in the epidermal growth factor receptor (EGFR) gene, replacing standard chemotherapy.

A study in British Columbia demonstrated that the survival of advanced lung cancer patients significantly improved after the introduction of targeted therapies in 2010 and immunotherapies in 2017.[34] In general, an increase in 1-year or 5-year survival from lung cancer has been documented over different periods between 1992 and 2017 in Canada, both overall and across cancer stages.[15, 35, 36] These results are promising and indicate that the prognosis of lung cancer patients diagnosed at advanced stages is expected to improve in the coming years as genetic testing of tumour samples becomes more readily available, newer genetic mutations are identified, and newer targeted therapies are introduced to treatment regimens.[37]

Decision-makers in Quebec are particularly interested in evaluating the clinical trajectories of advanced lung cancer patients receiving costly treatments, such as targeted therapies aimed at improving prognosis, with a focus on survival, treatment adoption, and timeliness.[38] One priority of the Programme québécois de cancérologie, as outlined in its 2024-2026 action plan, is

the use of effective, relevant, and innovative treatments, achieved through several objectives, including supporting the structured introduction of therapeutic innovations, measuring their benefits in real-world care, and ensuring equitable and timely access to cancer treatments.[39] Quebec has one of the highest lung cancer incidence rates in Canada (81.5 cases per 100,000) and an estimated 9,900 cases projected for 2024.[40] With approximately 70% of lung cancers in Canada diagnosed at advanced stages, novel therapies to extend survival are particularly critical for improving outcomes in this high-burden population.

2.2. Social determinants of health

According to the World Health Organization's (WHO) framework on social determinants of health, specific structural components of a society can affect individuals' health and contribute to health inequities.[41] That is, upstream societal system factors like economic, social, and political policies, along with social and cultural norms, shape the distribution of societal resources, rights, and powers across individuals, including systems of prestige and discrimination, thereby creating social hierarchies.[41, 42] This process of social stratification further impacts intermediary (i.e., downstream) social determinants of health, which are the "conditions in which people are born, grow, work, live, and age" and represent material circumstances, behaviours, and biological and psychosocial factors that directly impact individuals' well-being.[41, 43] The health care system itself is an intermediary social determinant that has an important role in health and health inequities. It acts on differential vulnerability or exposure to health-compromising factors and differential consequences of ill health caused by social stratification.[41]

Socioeconomic position, social position, or social identity represents the location of individuals within social hierarchies.[41, 44] Socioeconomic position is mainly measured through education, occupation, and income, the three main dimensions of SES, which is a measure of an individual's human capital and financial resources.[45] Each of the three dimensions captures different types of resources (knowledge, prestige, power, and money) and is a proxy for measuring social position.[41, 45] Sociodemographic and geographic characteristics - such as age, gender, race/ethnicity, immigration status, and place of residence - are also considered meaningful measures or indicators of social identity (i.e., equity stratifiers) in Canada.[3, 46, 47, 48] For example, social identity related to race or gender can lead to differential educational, employment,

and income opportunities. Additionally, many of these social identities are associated with experiences of discrimination, whether systemic or interpersonal.[41, 49]

Social position or identity (hereafter referred to as social status) is a structural determinant of health because it shapes individuals' vulnerability to and exposure to harmful risk factors beyond their personal control. Understanding the relationship between social status and health outcomes is essential for identifying health inequities—disparities that are unjust, avoidable, and remediable.[41]

2.3. Inequalities in lung cancer outcomes

Two primary approaches are used to study inequalities in cancer survival: analyzing the association between social status and survival or examining how social status influences intermediary social determinants of health, such as access to healthcare services or experiences within the healthcare system.[50] By studying inequalities in survival and healthcare, we can narrow down on individuals who may be experiencing survival inequities.

Many lung cancer studies have examined how social status influences cancer care or how either of these impact survival. However, most of the research on social statuses other than SES (e.g., race/ethnicity or immigrant status) has been primarily conducted in the United States or Europe.[51-65] In general, socially disadvantaged groups experience worse lung cancer care or survival, but not always (i.e., immigrant populations).

2.3.1. Inequalities in survival

2.3.2.1 Socioeconomic status

Education and income are two of the most common indicators of SES used in health inequality research.[41] Although these measures fall under the common SES umbrella, each measure is believed to impact health through different mechanisms. Researchers have examined individual-level SES and used area-level SES from census data as a proxy when individual data is unavailable.

Education

Education is typically measured in terms of the number of years of schooling completed or the academic milestones achieved (e.g., completion of secondary school). Education provides occupational opportunities with better income and improves literacy or factual knowledge necessary for disease prevention and management.[41, 45] Time spent in school also reduces time spent engaged in health-damaging activities. It promotes participation in cognitively challenging activities across one's lifespan, which can protect against dementia and increase survival.[45] Education can also impact the type of individuals included in our social network in terms of transferable knowledge and resources that they possess.[45] An advantage of using education as a measure of SES among older adults is that it is not affected by other changes later in life, like retirement. Therefore, it may be more closely related to lifelong accumulated wealth like home ownership.

Many studies have investigated the association between individual-level measures of education and survival,[66-84] while few have also assessed area-level measures.[59, 85-87] Among all identified studies on individual-level education, less than half reported a positive association between higher levels of education and survival.[70-73, 75, 76, 80, 81, 84] Some observed no association in the overall study samples but reported a protective association within categories of sex.[66-69] One study reported an inverse relationship,[77] and other studies found no association.[74, 76, 78, 79, 82, 83, 88] All studies on area-level measures of education reported a positive association with survival. Overall, the evidence on inequalities in lung cancer survival according to education level is mixed, partly due to heterogeneity in studies with respect to populations included, outcomes assessed, and covariate adjustments. For example, two studies included only lung cancer patients who received surgery and assessed survival from the time of surgery instead of diagnosis.[86, 87] Similarly, one study was conducted among non-smokers [77] and another among early-stage lung cancer patients.[85] Most studies estimated hazards of death,[59, 67, 69, 71-74, 76-83, 85, 86, 88] while others reported 30-day, 1-year, or 5-year survival probabilities or relative excess risks. About half of all studies (n=11) provided estimates adjusted for stage at diagnosis, a potential mediator of SES-based inequalities in lung cancer survival. A systematic review that included the majority of the studies on education found that some also

adjusted for smoking and treatment, which are also potential mediators.[89] Most studies on individual-level education were conducted in European countries, where national health registries often include education data. In contrast, studies on area-level education were undertaken primarily in the United States since national registries lack individual-level measures. Only two Canadian studies examined individual-level education and lung cancer survival. One study, with 225 patients, had inconclusive results.[83] The other found a protective effect of education but did not account for sex differences.[84]

Income

Income represents the most direct measure of material resources an individual possesses. Although income is primarily derived from paid employment, it also encompasses other sources of earnings and losses, including childcare benefits, social assistance or welfare support, tax deductions or credits, pensions, rental income, investment earnings, and more. Income can be calculated in many ways. Gross income is obtained by adding all sources of earnings within a fiscal year, while net income subtracts losses (e.g., taxes paid) from gross income. Household-level annual income can also be calculated by summing the net income of each household member (i.e., family income), which can further be adjusted for household size (i.e., equivalized household income) to account for economies of scale.[41] The advantage of using household net income over other income measures is that it accounts for disposable income to all family members, including children and those who are not primary income earners.

Income facilitates access to goods and services, including scarce material resources, that may be critical for maintaining good health, such as housing, food, transportation, and health care services and treatments.[41, 45] Low income can also negatively impact health through stress induced by social comparisons, reduced social resources, and low public trust and participation.[41] In turn, poor health can reduce income earnings, limiting access to resources and perpetuating poor health.

In contrast to education, the impact of income on lung cancer survival has been extensively studied, with most of the research focusing on area-level rather than individual or household-level measures of income. Of the 50 studies identified through a systematic review and independent database searches, 13 focus on individual-level income and 37 on area-level income.[69, 84, 88-98] Similar to the studies on education, there is variation between studies regarding survival outcomes and

covariate adjustment.[99] There is also variation in how income is measured and aggregated.[99] For example, for areas with an average population varying from 50 to 90,000 residents,[89] income has been summarized as median household income (gross or disposable),[86, 87, 93-96, 98, 100-111] percent of households below the poverty line,[61, 90, 106, 112-115, 116 , 117] or the income domain of a deprivation index.[92, 118-121] Overall, in a meta-analysis of seven studies on individual-level measures of income, lung cancer patients with lower income had a hazard of dying that was 1.13 times the hazard of those with higher income (95% confidence interval 1.08-1.19 and I-squared=0%).[89] No meta-analysis was conducted for studies on aggregate measures of income; however, most studies observed a trend similar to that for individual-level income.[89, 99] Of the 11 studies conducted in Canada that incorporated area-level measures of income,[93-96, 98, 100-105, 114] nine studies reported that higher income is associated with better lung cancer survival,[93-96, 100-105] one study found that lower income is associated with better survival,[98] and another reported no significant association between income and survival.[114] The two Canadian studies that incorporated individual-level income found a similar trend.[21, 84] However, one study did not account for sex differences and assessed income quintiles without adjusting for regional variations in income by province or rurality.[84] The other study also did not adjust for sex differences and reported stage-stratified estimates only.[21]

Deprivation index

Individual-level SES data is often unavailable for health inequality research, leading researchers to rely on area-level SES measures as proxies. For this purpose, individual-level SES information from the census is aggregated over the smallest possible geographic units to minimize measurement error when linking individuals to areas.[50] Additionally, rather than focusing on a single aggregated SES indicator (e.g., area-level income, education, or occupation), composite measures combining multiple SES domains (i.e., deprivation indices) are often used. These deprivation indices account for the high correlation among different SES indicators, and their application prevents from making inferences regarding one SES domain when, in fact, multiple markers of socioeconomic deprivation within one area may be at play. There are two main categories of indices: material deprivation index and social deprivation index. Material deprivation indices are based on dimensions of SES directly linked to access to material resources (e.g., employment, income, education), whereas social deprivation indices are based on SES domains

related to social support (e.g., living alone).[50] Area-level SES measures also account for context of place, so it may be challenging to disentangle the aggregate effect due to the SES of individuals living in an area versus an area's ecological effect.[50] For example, most materially deprived regions of Canada may have lower levels of health-promoting factors like walkability, clean air, or green spaces,[122] and reduced access to health care services.[123] In contrast, the advantage of using area-level SES deprivation to study health inequalities is the identification of neighbourhoods that may be eligible for interventions that simultaneously address barriers emanating from individual-level SES and neighbourhood-level deficiencies.[124]

At least 49 studies have examined the relationship between area-level SES indices and lung cancer survival.[85, 89, 125-130] Similar to studies on area-level income, there is substantial variation in the types of indices analyzed, the geographic population sizes for which aggregated SES measures are used, and the covariates included in each analysis.[89, 99] Despite this heterogeneity, all studies consistently reported an inverse relationship between area-level deprivation and lung cancer survival, as indicated by hazard ratios, median survival times, survival rate ratios, or survival rate differences. No study in Canada has examined the association between area-level deprivation indices and lung cancer survival in Canada.

2.3.2.2 Sociodemographic status

Age and gender

Lung cancer survival decreases with age at diagnosis and is lower among men relative to women.[131] For example, in Canada, the 1-year survival probability for individuals who are 55-64, 65-74, and 75-84 years is 47%, 40%, and 30%, respectively.[21] Similarly, the 1-year survival probability for women is 49% and for men is 40%.[21] The impact of age and sex on cancer prognosis is often attributed to biological differences.[132] For example, ageing increases the prevalence of comorbidities and frailty, both of which can affect treatment efficacy and survival.[133] Similarly, genetic and hormonal differences between men and women can influence tumour biology or immune responses that ultimately impact cancer progression.[133]

However, the influence of age and sex on cancer survival may also reflect broader social and systemic issues. Ageism, a process of stereotyping, prejudicing, or discriminating against individuals based on their age, negatively impacts health outcomes.[134, 135] In Canada, seniors

(65 years and older) are more vulnerable to discrimination and social isolation.[3] The COVID-19 pandemic further exposed systemic ageism within the healthcare system, such as inequities faced by older adults in long-term care and cancer care.[136-138] One Canadian study reported that older lung cancer patients are more likely to incur diagnostic and treatment delays.[139] Older individuals are generally underrepresented in clinical trials for cancer treatments, including trials on lung cancer therapies.[140-142] A lack of evidence on the efficacy of treatments in older patients can further limit them from receiving appropriate care.[143] As Canada's population ages, there is increasing concern about age-based inequalities along the cancer care continuum, which may negatively impact survival outcomes.[144]

Gender is a social construct shaped by cultural and societal expectations linked to an individual's sex. As a result, when gender data is lacking in health disparity research, sex is often used as a proxy measure, even though it may not fully capture the nuances of gender identity and its associated health impacts. In general, women are known to be less represented in clinical trials and, therefore, experience the same treatment struggles as older patients, emanating from clinical uncertainties in treatment efficacy and toxicity.[145] They are also more likely to be underrepresented in the studies informing lung cancer screening guidelines.[146-148] In contrast, women are also more likely to be diagnosed at earlier stages and receive surgery.[148-151] Among several explanations, it has been proposed that women's earlier diagnoses may be due to more timely contact with the healthcare system.[149, 150] In Canada, women were more likely to self-present for participation in a pilot lung cancer screening program.[22]

Overall, age- and gender-based inequalities in lung cancer outcomes are understudied, including in Canada.[146, 152-154] Furthermore, the relationship between these factors and lung cancer survival is not straightforward as it may vary by individuals' SES, race/ethnicity, sexual orientation, physical status, indigenous ancestry, nationality, or place of residence.[146, 152, 153, 155, 156]

Immigrant status

Non-immigrants are individuals who are born in Canada. In contrast, immigrants are individuals admitted to Canada as permanent residents (i.e., landed immigrants) for economic or family reasons or refuge. Individuals also arrive as non-permanent residents, including refugee claimants

and those holding work or study permits. Those who are or have been permanent residents represent 23% of the Canadian population, of which one-third have been non-permanent residents prior to becoming immigrants.[157] Over time, the birth countries of immigrants to Canada have shifted from primarily Europe to Asia and Africa, contributing to the country's growing linguistic and ethnocultural diversity.[157] Furthermore, of 19% of individuals in Canada who are non-Aboriginal and non-White in colour or non-Caucasian (i.e., visible minorities), two-thirds are immigrants.[158]

Canadian immigration policies, such as the point-based ranking system, are designed to promote the selection of immigrants who will contribute to the country's economic growth.[159] Consequently, more than half of all immigrants admitted to Canada fall under the economic immigrant category.[157] Furthermore, the selection process requires immigrants to undergo medical screening tests to prevent a high burden on the healthcare system.[159] These policies, in addition to positive self-selection and healthier habits, may explain why immigrants tend to be healthier than non-immigrants upon arrival in Canada.[159] However, this healthy immigrant effect tends to dissipate over time, potentially due to an acculturation to the Canadian lifestyle, which involves health behaviours that are assumed to be riskier and a poorer diet, a lack of linguistic assimilation, a lack of access to healthcare, and/or discrimination (interpersonal or institutional).[160] Immigration is considered a social determinant of health, as immigrants face social exclusion driven by discriminatory workforce policies and attitudes.[161] Additionally, cultural, linguistic, and informational barriers and discrimination within the healthcare system hinder their access to care and health.[162]

Cancer is a disease that impacts individuals at older ages. Therefore, it is possible that the healthy immigrant effect is not observed among immigrants with cancer. In general, there is little information on how cancer survival varies by immigrant status due to the lack of this information in health administrative data. Previous studies in the United States (US) show mixed cancer survival outcomes for foreign-born individuals compared to their US-born counterparts. Some studies indicate worse survival among foreign-born individuals with certain cancers, such as liver and cervical cancers,[163-167] while other research suggests better outcomes, including for lung cancer.[168-173] However, these findings may not be generalizable to Canada due to differences in immigrant composition and healthcare systems.

One study that combined all cancers in Ontario found that recent and non-recent immigrants had better survival outcomes than non-immigrants; however, this survival advantage decreased with time since immigration.[174] Another study conducted in Ontario, specifically on lung cancer survival, found a similar trend. However, 75% of immigrants in this study had been in Canada for less than 24 years, and at least 50% belonged to the sponsored family class.[175] Given that the median age at lung cancer diagnosis is approximately 70 years, this means that most immigrants in this study arrived in Canada at the age of 46 years or after, which accounts for only 15-16% of all immigrants.[157] Finally, one study conducted at the national level with census data found higher 5-year survival probabilities in immigrants relative to non-immigrants. However, the analysis included prevalent cases of lung cancer and did not account for sex-based differences by immigration status.[84]

2.3.2.3 Place of residence

Location of residence can impact individuals' socio-economic status. For example, individuals living in rural areas compared to urban areas in Canada often achieve lower levels of post-secondary education.[176] This further limits them in skills essential for employment in an economy transitioning from resource-based to knowledge-based industries.[176] Overall, rural areas have slower economic development with decreased employment opportunities and lower wages.[176] An individual's place of residence also influences their exposure to health-promoting or harmful factors, as well as the healthcare resources in proximity at their disposal, including healthcare. Geography-based health inequalities arise from the unequal distribution of these elements (SES, physical environment, and access to healthcare) across different spatial scales: countries, provinces, cities, or neighbourhoods.[177] In Canada, the focus has been on rural health inequality as rural area residents have been shown to have less healthy habits (e.g., smoking and obesity) and a shorter life expectancy than urban residents.[178-180] Rural areas, relative to urban areas, experience higher mortality from low SES, injuries, suicide, cardiovascular disease, and diabetes.[180] In contrast, rural areas experience lower mortality from cancer,[180] which may be due to lower incidence rates of cancer and does not necessarily translate into better survival relative to urban areas. A meta-analysis of studies on cancer survival in developed countries, including studies on lung cancer survival, suggests that cancer survival is lower in rural versus urban residents.[181]

Studies on cancer survival have measured rurality in various ways, including accessibility or remoteness indices, county or provincial classifications, population density, commuting patterns, metropolitan versus non-metropolitan influence, and distance from essential services or healthcare centers. Rurality is better conceptualized as existing on a continuum rather than a binary distinction.[181] A national-level study in Canada examined 3-year lung cancer survival between urban and rural areas and found minimal survival differences (26% in rural areas versus 29% in urban areas).[182] However, this analysis did not account for sex differences, even though the incidence of lung cancer in rural areas is higher among men,[183] who are known to have lower lung cancer survival rates in comparison to women. Rurality was also assessed as a dichotomous measure; however, most studies that used non-binary classification of rurality have shown that cancer survival is worse in most remote patients.[181] An Ontario study on lung cancer survival found that patients living in areas with an Ontario rurality index value indicative of small urban areas had slightly lower survival rates than those in large urban areas (hazard ratio: 1.05, 95% confidence interval: 1.03 to 1.08).[98] However, patients from all rural areas were grouped, and no significant difference in survival was found between rural and large urban areas. Interestingly, the study also reported lower survival for patients living near small hospitals than those living near teaching hospitals.

2.3.2. Inequalities in cancer care

The healthcare system is an intermediate social determinant of health that plays an important role in mediating and intervening on health inequities.[41, 184] To better understand inequities in cancer survival, most studies focus on inequalities along the cancer care continuum, particularly in diagnostic procedures and treatments, as they are avoidable or remediable.[154] Delays in diagnosis and treatment are also recognized as key factors driving SES-based inequalities in cancer survival.[185]

2.3.2.1 Diagnosis

Timely access to diagnostic procedures, whether influenced by the patient or healthcare provider, directly impacts the stage at which lung cancer is diagnosed.[50, 186] Delays in the diagnosis of lung cancer in the Canadian context have been associated with a lack of symptom recognition and complex diagnostic processes.[187] Previous studies have shown that socially disadvantaged

patients are less likely to use preventive cancer screening services [188-192] and less likely to receive diagnostic imaging for lung cancer.[95] In Canada, some barriers for low-SES individuals in accessing lung cancer diagnostic procedures include transportation costs, inflexible work hours, unstable housing, or limited access to emails.[193] The hypothesis that advanced stage at diagnosis partly mediates social and geographic inequalities in lung cancer survival has been reiterated multiple times in the literature.[194, 195]

A systematic review of factors explaining SES-based inequalities in cancer survival reported mixed findings regarding the mediating role of stage a diagnosis.[185] Another review on lung cancer survival found no differences in stage at diagnosis by SES; however, most studies in the review used area-level SES measures.[196] In contrast, a recent study in the United States found differences in stage at diagnosis for lung cancer by neighbourhood-level income and education.[197] A separate review also observed slight reductions in the effects of low income on lung cancer survival in studies that adjusted for stage at diagnosis compared to those that did not, but only when individual-level SES was considered; no such reductions were found for area-level SES.[89] One study on lung cancer conducted in Canada reported differences in stage at diagnosis by individual-level income,[97] whereas two other studies found no association between area-level income and stage at diagnosis.[101, 198] Other studies have also shown differences in the stage of lung cancer diagnosis by age,[199-201] sex,[148-151] and urban-rural residence.[202] One study found no association between immigrant status and lung cancer stage in Canada.[198] However, it is possible that differences in stage by immigrant status were masked in this study, as it was based on an analysis that combined individuals who arrived in Canada more than 33 years ago (long-term immigrants) with non-immigrants.

Overall, lung cancer patients have a poorer prognosis than those with other major cancers, even at stage I. As a result, stage at diagnosis is believed to have a larger role in driving SES-based inequalities in survival for cancers that have better prognoses than for lung cancer.[50, 89]

2.3.2.2 Treatment

SES-based differences in lung cancer survival are smaller than for other cancers, possibly because the majority of individuals are diagnosed at advanced stages when treatment effects are limited.[89, 101] Meta-analyses indicate that treatments like surgery and chemotherapy may

partially explain SES-based inequalities in lung cancer survival, whereas radiotherapy does not appear to play a significant role.[203] Chemotherapy is used across both early and advanced stages of lung cancer. However, the stronger association of SES with surgery—primarily performed in the early stages of lung cancer—and the lack of association with radiotherapy—more commonly prescribed in advanced stages—suggest that SES-based inequalities in lung cancer survival are driven by differences in treatment (surgery and chemotherapy) used in early stages. Therefore, SES-based inequalities in survival are more likely to be observed at early stages of lung cancer rather than at late stages. The difference in 3-year survival probabilities between the lowest and highest family income quintiles among lung cancer patients in Canada is larger for stage I (high: 84%, low: 73%) versus stage IV (high: 8%, low: 6%).[21]

Although breakthrough therapies for advanced lung cancer seem promising for improving population-level survival, not all patients with advanced lung cancer may experience survival benefits due to inequities in drug access by place of residence. This can be seen with targeted therapies. In 2009, Health Canada approved gefitinib as the first targeted therapy for lung cancer patients with tumours harboring mutations in the EGFR gene.[204] However, testing for genetic mutations only became widely available in Canada in March 2010 for one year through a compassionate use program for gefitinib funded by the pharmaceutical industry.[204] At that time, EGFR tests were performed only in large cities or academic centers, limiting access in rural areas.[204-206] Subsequently, access to EGFR mutation testing and EGFR-targeted drugs varied by province. For instance, testing received provincial funding in April 2011 in British Columbia and September 2014 in Ontario.[204, 207, 208] Similarly, gefitinib was covered through public drug insurance programs in October 2010 in British Columbia and November 2011 in Quebec.[34, 38] Furthermore, regional variation exists within provinces in the implementation of reflex genetic testing.[208] Reflex testing, the automatic genetic analysis of tumour samples conducted by pathology labs without requiring additional requisitions, can impact access to timely and optimal treatment.[209]

Access to targeted therapies may also vary according to individuals' socioeconomic and sociodemographic status. Intravenous chemotherapies administered in-hospital fall under Canadian provinces' universal public coverage. In contrast, newer oral therapies are filled in community pharmacies and taken at home and are, therefore, not systematically covered by

provincial public health insurance programs. Patients' out-of-pocket costs thus vary by province, age, income, and type of drug plan (if any). In western provinces, like British Columbia, targeted therapies are publicly funded for all cancer patients who meet treatment indication criteria.[210, 211] In contrast, Ontario publicly funds these drugs through an exceptional access program, which only covers all individuals aged 65 or over, those under 25 without private insurance, or those on social assistance.[210] In Quebec, public coverage is available for all individuals aged 65 or over, those on social assistance, and those under 65 who do not have private insurance. Both public and private plans involve varying out-of-pocket costs due to co-payments and deductibles that can act as financial barriers.[212-216] In Canada, smoking status, age, gender, income, education, and ethnicity have been cited as barriers to accessing biomarker testing and, hence, targeted treatments.[204, 205, 208, 209, 217] Overall, with the introduction of next-generation treatments prolonging the survival of advanced lung cancer patients, social and geographic inequalities in survival are anticipated to become more important due to pre-existing structural barriers and fragmented access to newer therapies.

2.4. Intersectionality

The underpinning of intersectionality theory is that different structures of societal oppression and privilege do not act independently; rather, their interconnectedness creates unique lived experiences, which are fluid with space and time.[218] This theory was developed in the late 1970s by black feminists. Kimberlé Williams Crenshaw and Patricia Hill Collins, critical race and social theorists, respectively, further advanced and defined this theory as "intersectionality."[156, 218] Like the social determinants of health framework, intersectionality theory adopts a social justice approach to addressing health inequalities. However, it evaluates individuals' social locations based on intersecting social categories rather than single-axis categories that assume uniform experiences of power, privilege, and discrimination.[156] As a critical analytical lens, intersectionality serves as a catalyst for transformative change.[218]

The concept of intersectionality is increasingly being highlighted in the Canadian literature on inequalities in cancer outcomes,[153, 184, 219-221] including studies on lung cancer survival.[84, 175] Cancer patients facing the compounding effects of multiple social disadvantages—such as racism, discrimination, poverty, and ableism—are characterized as marginalized populations who

may experience wider inequities in health and care across the cancer continuum.[184] These inequities manifest as higher rates of advanced cancer diagnoses, increased mortality, and limited access to timely, guideline-concordant treatments.[222, 223]

To our knowledge, no studies have specifically examined inequalities in lung cancer survival across individuals with intersecting social statuses. However, findings from some prior research suggest the importance of considering intersectionality when investigating inequalities in lung cancer care and outcomes. For example, one systematic review has shown that age-based disparities in lung cancer survival are larger in women relative to men.[152] In Canada, the effect of immigrant status on lung cancer survival has been shown to decrease with older age and sex.[175] In another study in the United States, the effect of SES on lung cancer survival was more substantial among non-Hispanic Black patients in comparison to non-Hispanic White patients.[224] Similarly, another study found that Black patients were significantly more likely than White patients with similar profiles of SES to be diagnosed at advanced stages of lung cancer.[197]

Overall, the integration of intersectionality into quantitative analyses of inequalities along the cancer care continuum remains relatively rare in Canada and the United States.[218, 225, 226] This gap is partly due to past limitations in the availability of datasets with sufficient observations and robust quantitative methodologies needed for this type of work.[225, 227, 228]

2.5. Summary of evidence and knowledge gaps

Lung cancer remains one of the most common and deadliest cancers in Canada, with mortality rates surpassing those of colorectal, breast, and prostate cancers. This high mortality stems from the fact that most cases are diagnosed at advanced stages, when treatment options are limited, largely due to the historical absence of targeted, organized screening programs. With the anticipated rollout of provincial screening programs across Canada and advances in precision medicine, inequalities in survival may emerge or intensify. Preventing inequities when planning lung cancer care in Canada first requires understanding whether socially vulnerable populations face unequal health outcomes and identifying contributing factors within cancer care systems.

Although there is a large body of evidence on the variation of lung cancer survival by income, less is known about its relationship with education, immigration status, and place of residence. In Canada, most studies have evaluated socioeconomic factors at the area level or have been conducted within specific provinces. The very few Canadian studies that included individual-level measures were conducted at a provincial level and/or had methodological limitations. Studies have also shown that cancer patients living in the most remote rural areas have worse survival. Yet, previous studies on lung cancer survival have mainly examined rural areas as a unit, which may be an oversimplification. Most importantly, there are uncertainties regarding the role of stage at diagnosis in mediating existing inequalities, and formal mediation analysis has yet to be applied to quantify its role.

Differences in curative treatment received for early-stage lung cancer may partly drive SES-based inequalities in survival. However, next-generation therapies are changing the paradigm of advanced lung cancer. These treatments may amplify existing inequalities in lung cancer survival, given that the majority of lung cancer patients are diagnosed at an advanced stage. Yet, not much is known about social and geographic inequalities in the survival and treatment of advanced lung cancer patients in the era of breakthrough therapies.

Finally, examining health inequalities through single markers of social determinants risks oversimplifying the lived experiences of populations at the intersection of multiple social locations. To advance intersectional research and inform policies and interventions, it is essential to map inequalities across populations while accounting for interlocking systems of oppression and privilege. While intersectionality-informed quantitative analysis can uncover the complexities and nuances of individuals' social contexts, its application in cancer research remains limited.

Chapter 3. Methods overview

3.1. Data sources

The studies included in this thesis relied on multiple data sources. In this chapter, I briefly describe each dataset, with further details available in the manuscripts and their supplementary materials.

3.1.1. Data linkage with the Canadian Community Health Survey

For Manuscript 1, I was granted access by Statistics Canada to a data linkage project titled *'Linkage of the Canadian Community Health Survey to mortality, cancer, hospital administrative files, Census (short-form), and tax data'* through the Quebec Inter-University Centre for Social Statistics (QICSS) McGill-Concordia lab.[229] This university-based Research Data Centre (RDC) provided a secure environment to analyze anonymized and de-identified individual-level data. The linkage was done by Statistics Canada in the Social Data Linkage Environment (SDLE), a secure environment for linking datasets, to examine the impact of social determinants of health on population-level health outcomes.

All datasets in this linkage project have been linked to the Derived Record Repository (DRD), which is housed within the SDLE and contains unique identifiers for the population found in Statistic Canada source files. The records of the datasets used in Manuscript 1 (see below) were probabilistically linked to the DRD through non-unique identifiers such as age, sex, address, and date of birth. Each record was then assigned a unique anonymized record ID stored in a key registry, which was used for linking different datasets together. The Canadian Community Health Survey (CCHS, see below) data served as the cohort base for this linkage project. That is, unique anonymized record IDs in the DRD belonging to CCHS respondents were used to extract matching records from other databases that were independently linked to the DRD.

Canadian Community Health Survey (CCHS): 2000-2017

The Canadian Community Health Survey (CCHS) is a recurring national survey conducted by Statistics Canada that collects self-reported information on SES (e.g., education), sociodemographic status (e.g., immigration status), health status, healthcare utilization, and health behaviours among individuals aged 12 and older, excluding specific populations such as Canadian

Forces members, residents of special care facilities, prisons, First Nations reserves, Crown lands, and some remote regions. All CCHS survey waves (2000/01, 2003/04, 2005/06, and 2007 to 2017) included in the linkage project were used for the study in Manuscript 1.

Each CCHS record in the linkage project was assigned an individual-level weight that accounts for the survey design, non-response, individuals who did not consent to share their survey data with Statistics Canada partners or link it with other data, and those whose records could not be linked to the DRD. Each record was also accompanied by 500-1000 bootstrap weights for variance estimation.

Canadian Cancer Registry (CCR): 2000-2016

The Canadian Cancer Registry (CCR) is a Statistics Canada database that integrates data from provincial and territorial cancer registries, providing standardized information on all cancer diagnoses in Canada since 1992. Staging data has been systematically included in the CCR since 2010. However, cancer incidence data from Quebec from 2010 and onward was missing from the CCR, since it had not submitted at the time of linkage.

The linkage project incorporated data from the CCR tabulation master files (TMF) from 1992 to 2016. However, for Manuscript 1, only data from individuals diagnosed with lung cancer between 2010 and 2016 were used.

Canadian Vital Statistics Database (CVSD): 2010-2017

The Canadian Vital Statistics Database (CVSD) is a Statistics Canada dataset that compiles data on all deaths in Canada through information provided by provincial and territorial vital statistics registries. It includes details on individuals' demographics, as well as their dates and causes of death. CVSD data from 2000 to 2017 were sourced for the linkage project, but only data from 2010 to 2017 were used for Manuscript 1.

Annual Income Estimates for Census Families and Individuals (T1 Family File or T1FF): 2005-2015

The Annual Income Estimates for Census Families and Individuals (T1 Family File or T1FF), compiled by the Canada Revenue Agency (CRA) and provided to Statistics Canada, include data from income tax returns and the Canada Child Tax Benefit (CCTB). It contains information on

sources of income and demographic indicators for both tax filers and non-filers. T1FF data from 1993 to 2017 were linked to the CCHS; however, only data from 2005 to 2015 were used for Manuscript 1. The primary information extracted from the T1FF was after-tax family income.

Postal Code Conversion File – Plus (PCCF+)

The Postal Code Conversion File – Plus (PCCF+) is a SAS program that links 6-digit postal codes in Canada to standard geographic areas defined by Statistics Canada (e.g., dissemination areas), which are used to produce census data and other statistics, such as neighbourhood-level socioeconomic measures.[230] The primary purpose of using PCCF+ in this thesis was to derive urban-rural measures for place of residence by linking individuals' postal codes to census-defined geographic areas. There are different versions of the PCCF+ to account for changes in geographic areas and statistics by census year. Since each postal code can be associated with multiple geographic areas, the PCCF+ program uses census-derived weights for probabilistic allocation, providing a more accurate geographic assignment, particularly for rural postal codes linked to multiple areas. In contrast, the PCCF file assigns postal codes to a single geographic area based on the highest concentration of dwellings but is less accurate, particularly for rural areas,[231] and was therefore not used.

The PCCF+ program was needed to convert postal codes recorded at the time of diagnosis in the CCR and at the time when income was measured (e.g., two years before diagnosis) in the T1FF. However, the PCCF+ program was not part of the linkage, and different versions of the program were added to my account at the RDC through the Data Liberation Initiative. The Data Liberation Initiative is a membership made available to post-secondary Canadian institutions by Statistics Canada to improve access to data and make it affordable.[232] The versions of the PCCF+ that were used depended on the year in which the postal code was recorded:

PCCF+ version	Census year	Year of postal code
5k	2006	2008-2009
6d	2011	2010-2014
7e	2016	2015-2016

3.1.2. Quebec's health administrative data

For Manuscript 2, I used Quebec's health administrative data, accessed through a Health System Impact Fellowship completed within the cancer unit at INESSS. This linked data is unique as it includes information on specific treatments, unlike the other linked datasets used in this thesis that do not include treatment information. The specific databases used for Manuscript 2 (detailed below) are managed by the Régie de l'assurance maladie du Québec (RAMQ) and the Ministère de la Santé et des Services sociaux (MSSS). They are made available to INESSS in de-identified form through a tripartite agreement between MSSS, RAMQ, and INESSS.[233] More information on these databases can also be obtained from previous INESSS reports.[234, 235]

Pharmaceutical services (Services pharmaceutiques, SMED): 2001-2019

SMED is a database managed by the RAMQ that contains information on all drugs dispensed in community pharmacies and reimbursed through the public drug insurance plan. Between 2001 and 2019, users of targeted therapies were identified in this database through nonproprietary names (INN) and drug identification numbers (DIN). These users served as the cohort base, and their records were linked to information from other databases (see below).

Health insurance registry (Fichier d'inscription des personnes assurées, FIPA): 2001-2020

FIPA, managed by RAMQ, includes data on all individuals covered by public health insurance and prescription drug plans. It provides information such as sex, date of birth, date of death, and the duration of drug insurance coverage. RAMQ has linked full postal codes (which INESSS cannot access) to geographic descriptors from the 2016 census using the PCCF file (not PCCF+) with the single link method. As a result, FIPA also includes data on income, education, and immigration status at the dissemination area level, representing approximately 400 to 700 individuals.

Hospitalization (Maintenance et exploitation des données pour l'étude de la clientèle hospitalière, MED-ECHO) and fee-for-service physician billing (Services rémunérés à l'acte, SMOD): 2001-2019

MED-ECHO, managed by MSSS, contains data on hospitalizations and day surgeries, while SMOD, managed by RAMQ, provides information on physician services reimbursed on a fee-for-

service basis. Both databases include records of medical procedures: MED-ECHO identifies procedures using codes such as the Canadian Classification of Health Interventions (CCI) and the Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures (CCP),[236, 237] while SMOD uses physician billing codes. For Manuscript 2, these databases were used to identify biopsy procedures with CCI, CCP, and physician billing codes, all validated by medical experts.

3.1.3. Data linkage with the Canadian Cancer Registry

For Manuscript 3, I was granted access by Statistics Canada to a data linkage project titled '*Linkage of Socioeconomic and Treatment Data to the Canadian Cancer Registry*', also known as the Canadian Partnership Against Cancer (CPAC) linkage project. Similar to the data used for Manuscript 1, the data from this linkage project was accessed through the QICSS McGill-Concordia lab. The linkage was conducted in the SDLE, and the Canadian Cancer Registry (CCR) served as the cohort base for the linkage. Most of the databases linked to the CCR were the same as those linked to the CCHS in the previous project used for Manuscript 1, except for some subtleties listed below.

Canadian Cancer Registry (CCR): 2002-2015

Similar to the previous linkage, this second linkage project included all cancers identified in the CCR TMF from 1992 to 2015. However, for the study in Manuscript 3, only lung cancer cases identified between 2002 and 2013 were retained.

Canadian Vital Statistics Database (CVSD) and T1 Personal Master File (T1PMF-D): 2002-2014

Unlike the previous linkage, where death information was obtained directly from the CVSD, in this linkage project, death information was ascertained from three sources: the CVSD, the T1 Personal Master File – Death subset (T1PMF-D), and the CCR. The T1PMF is a dataset compiled by the Canada Revenue Agency (CRA). It includes data on all individuals who filed income tax returns or received Canada Child Tax Benefits (CCTB), as well as information on their children and spouses. The T1PMF-D, a subset of the T1PMF, specifically contains death-related information extracted from tax files. While death information was available from 1992 to 2014, the study in Manuscript 3 used data from 2002 to 2014.

Annual Income Estimates for Census Families and Individuals (T1 Family File or T1FF): 1999-2015

The T1FF was used to obtain information on after-tax family income of individuals included in the study for Manuscript 3. While income information was available from 1992 to 2015, only data from 1999 to 2012 were used.

Longitudinal Immigration Database (IMDB): 1980-2015

The Longitudinal Immigration Database (IMDB) includes information on landed immigrants (permanent residents) and non-permanent residents who arrived in Canada since 1980. It belongs to Immigration, Refugees, and Citizenship Canada and is managed by Statistics Canada. The database includes information related to immigration and other data on sociodemographic characteristics and SES (derived from tax files). Although the linkage project included information on individuals immigrating to Canada between 1980 and 2015, only information on those who arrived between 1980 and 2013 was used in the study for Manuscript 3. The IMDB was used exclusively to identify individuals in the CCR who were classified as immigrants. No additional information from the IMDB was incorporated into the study.

Discharge Abstract Database (DAD) and National Ambulatory Care Reporting System (NACRS): 2004/05-2015/16

The Discharge Abstract Database (DAD) is a comprehensive administrative dataset that includes information on hospital discharges from acute care in all Canadian provinces and territories, except Quebec. In addition to administrative information and sociodemographic characteristics, DAD includes clinical information associated with hospitalization, such as interventions and diagnosis codes.

The National Ambulatory Care Reporting System (NACRS) is an administrative database that captures information on ambulatory care visits in hospital, outpatient, and community-based settings. It includes records on day surgeries, emergency department visits, diagnostic imaging, and oncology services, all of which are associated with intervention and diagnostic codes.

The primary purpose of using DAD and NACRS in the study for Manuscript 3 was to identify diagnosis codes that enabled the assessment of individuals' comorbidities. The linkage project

included data from DAD and NACRS covering the fiscal years 1994/1995 to 2015/2016 and 2002/2003 to 2015/2016, respectively. However, for the study in Manuscript 3, only records from January 1, 2005, to December 31, 2014, were used.

Postal Code Conversion File – Plus (PCCF+)

Like the linkage project used for Manuscript 1, the PCCF+ program was accessed through the Data Liberation Initiative and added to my account at the RDC. The PCCF+ program was only needed to convert postal codes recorded at the time of diagnosis in the CCR, and the versions of the program used depended on the year in which the postal code was recorded:

PCCF+ version	Census year	Year of postal code
4K	2001	2002-2004
5k	2006	2008-2009
6d	2011	2010-2014
7e	2016	2015-2016

In this linkage project, postal codes associated with the income records in the T1FF were not directly provided by Statistics Canada. Instead, Statistics Canada performed an in-house linkage of the postal codes using the PCCF+ programs to obtain census-specific geographic measures, which were then included in the T1FF. These geographic variables were derived from different census years, depending on the year of the T1FF record:

T1FF taxation year	Census year
1996-1999	1996
2000-2004	2001
2005-2009	2006
2010-2014	2011

3.2. Ethics

All data used for research presented in this thesis were secondary in nature. Ethics approval was obtained from McGill University's Institutional Review Board for Manuscripts 1 and 3 (A04-M30-22A) and Manuscript 2 (A04-M31-22A).

Chapter 4. Manuscript 1

4.1. Preface

This manuscript is the first in a series of three that collectively examine social inequalities in lung cancer survival across Canada, each addressing a distinct yet interconnected aspect of the issue. Here, I focus on individual-level social factors—income, education, immigration status, and place of residence—and their impact on survival, while also investigating the mediating role of advanced stage at diagnosis.

For this study, I leverage a unique data linkage project from Statistics Canada that integrates data from the CCHS, the CCR, income tax records, and other administrative datasets. This linkage provides individual-level measures of education, immigration status, and reliable (non-self-reported) income data, which is available close to the date of cancer diagnosis in the CCR. Made available to researchers in 2020 through RDCs, this resource provides as an innovative approach for investigating social inequalities in lung cancer outcomes. I was among the first researchers granted access to this dataset, enabling a timely and comprehensive analysis of lung cancer survival disparities in Canada.

The work in this manuscript establishes a foundation for subsequent manuscripts by demonstrating how inequalities in survival emerge and may persist beyond the mediating role of stage at diagnosis.

4.2. Social inequalities in survival among Canadian lung cancer patients and the mediating role of stage at diagnosis

Samia Qureshi[1*], Mireille E Schnitzer[2], Arijit Nandi[1,3], Erin Strumpf[1,4]

[1]Department of Epidemiology, Biostatistics and Occupational Health (EBOH), McGill University, H3A 1G1, Montreal, Quebec, Canada

[2] Faculty of Pharmacy and the Department of Social and Preventative Medicine, Université de Montréal, H3N 1X9, Montreal, Quebec, Canada

[3] Department of Equity, Ethics, and Policy, McGill University, H3A 1G1, Montreal, Quebec, Canada

[4] Department of Economics, McGill University, H3A 1G1, Montreal, Quebec, Canada

[*]**Corresponding Author:** Samia Qureshi

Address for Correspondence: samia.qureshi@mail.mcgill.ca

Key Words: lung cancer, stage at diagnosis, survival, health inequalities, socioeconomic status, income, education, immigrant status, urban, rural

Word Count: Abstract: 265**;** Main text: 3,531

Abstract

Income-based inequalities in lung cancer survival are well-documented in Canada; however, differences by other social factors and the contribution of advanced stage (IIIb/IV) at diagnosis remain unexplored.

Among respondents of the Canadian Community Health Survey (2000-2017) with a primary lung cancer diagnosis between 2010 and 2016 (N=4,430), we examined survival by income, education, immigration status, and an urban-rural spectrum for place of residence. Applying counterfactual-based mediation methods, we estimated the percentage of the total effect (TE) of each social factor on survival that is represented by the mediating effect of advanced stage at diagnosis (i.e., total natural indirect effect: TNIE).

We observed differences in weighted Kaplan-Meier survival curves across the urban-rural spectrum (logrank p-value: 0.0215). Although estimated with some imprecision, the largest hazard difference was between small and large urban cities (50.1 additional deaths per 10,000 person-months, 95% CI: -3.7 to 107.2). Point estimates suggest that 40% of this TE is mediated by advanced stage at diagnosis (TNIE: 20.0 deaths per 10,000 person-months, 95% CI: -11.1 to 47.5). Median survival difference for respondents with low versus high education was -3.5 months (95% CI: -5.0 to -0.2). Despite imprecision, the hazard difference estimates for income and education, also suggest a harmful association between low SES and survival, with corresponding TEs of 44.8 deaths (95% CI: -3.3 to 101.2) and 39.6 deaths (95% CI: -7.4 to 79.0) per 10,000 person-months, while smaller TNIEs were observed for both. No survival inequalities by immigration status were observed.

While lung cancer survival inequalities exist by individual-level SES, advanced stage at diagnosis may largely mediate inequalities found by urban city size.

Introduction

In Canada, half of all lung cancer patients are diagnosed at stage IV, which has a stark three-year survival probability of 5%.[1, 2] Socially and geographically disadvantaged patients may face poorer outcomes in part due to late-stage diagnosis,[3, 4] which is often an indicator of delayed access to care.

Income-based inequalities in lung cancer survival are well documented in Canada,[2, 5-12] but the impacts of other social determinants, like education, immigration, and area of residence, are less clear.[4, 13] Income is often considered a marker of SES tied to material resources, whereas education reflects knowledge and health literacy.[14, 15] In contrast, immigration status and place of residence may influence survival through language and cultural barriers,[16] discrimination,[14, 17] or limited physical access to care.[18] Information on individual-level SES or immigration is typically unavailable in administrative datasets, and area-level measures may not be ideal.[19] Recent studies have not been able to show survival differences between urban and rural areas,[2, 12] but this may be due to a dichotomous measure hiding nuanced differences.

A recent Canada-wide lung cancer study revealed inequalities in stage at diagnosis by individual-level income,[12, 13] reinforcing the need to assess its impact on survival inequalities. Reviews of studies on lung cancer populations outside of Canada suggest that stage at diagnosis is not associated with SES,[20] and report mixed findings regarding its role in mediating SES-based survival inequalities.[4, 21] However, most of these studies have either relied on area-based measures of SES or faced methodological limitations in investigating the mediating role of stage at diagnosis, such as using the difference method, which is prone to collider bias and assumes no interaction between the exposure and mediator.[21-23]

In this study, we evaluated lung cancer survival inequalities in Canada by individual-level income, education, and immigration status, and by place of residence on an urban-rural spectrum. For each social factor, we also quantified the mediating role of stage at diagnosis using causal mediation methods. Given that lung cancer progresses quickly and earlier diagnosis provides opportunities for curative-intent treatment,[24] identifying target populations for early diagnosis interventions is crucial for addressing social inequalities in survival.

Material and Methods

Study sample

A cohort study was conducted by selecting respondents from Statistics Canada's Canadian Community Health Survey (CCHS: 2000-2017) with a primary lung cancer diagnosis in the Canadian Cancer Registry (CCR: 2010 to 2016) (see details in Appendix A). Respondents diagnosed before 2010 could not be included due to non-systematic reporting of staging data in the CCR, and those diagnosed in Quebec could not be included because Quebec did not contribute to the CCR after 2010. Respondents with stage 0 (in-situ, pre-cancerous cells) or missing information on education or immigration status were excluded. Those missing information on stage or stage III subtype, income, date of birth, date of diagnosis, date of death after imputation were also excluded.

Data sources

In addition to the CCR incidence Tabulation Master File, CCHS data was linked to the Annual Income Estimates for Census Families and Individuals files, the Canadian Vital Statistics Database, and Postal Code Conversion File Plus (PCCF+) programs in SAS (see details in Appendix A and Table 4.3.1).[25]

Exposures

We used the adjusted after-tax family income from two years before the respondents' lung cancer diagnosis to avoid capturing disease-related income loss. This income was categorized as low, medium, or high based on population-level terciles defined by calendar year, province, and rurality (Appendix B). Education was grouped as follows: 1) less than secondary school completion, and 2) secondary school completion and/or post-secondary education. For immigration status, respondents were classified as non-immigrants (Canadian-born) or immigrants (foreign-born).

Urban-rural exposure was based on Statistical Area Classification (SAC) types derived from postal codes reported in the CCR at the time of diagnosis and PCCF+ programs in SAS (Appendix C). We consolidated SAC types into: 1) large cities (census metropolitan areas), 2) small cities (census agglomerations), 3) rural I (metropolitan influence zones with strong or moderate influence), and 4) rural II (metropolitan influence zones with weak or no influence, or territories).

Mediator

Cancer stage at diagnosis in the CCR was derived using the Collaborative Staging algorithm, which is based on the seventh edition of the American Joint Committee on Cancer's staging manual.[26] Stage I, II, or IIIa tumours were classified as early stages, while stage IIIb or IV tumours were classified as advanced stages due to their inoperability and lower survival rates.[27-29] Other potential mediators were not assessed in this study.

Outcome

Survival time was calculated in months from the date of diagnosis to death or administrative censoring (December 31, 2017), whichever came first. If only the day was missing for any date, it was imputed as the 15th of the month.

Covariates

We aimed to account for all confounding paths in our survival and mediation analyses. Based on our theoretical causal directed acyclic graph (DAG, Figure 4.2.1), the baseline confounders were age, sex, and year of diagnosis, which affect survival and are associated with exposures of interest either directly or indirectly through lung cancer incidence.[30, 31] Additionally, we accounted for province and rurality in analyses where income or education was the exposure, for province when place of residence was the exposure, and for race/ethnicity when immigration was the exposure. Due to multiple adjustments and limited cell counts, we could not adjust for race/ethnicity and immigration in our analysis of income or education, and for immigration in our analysis of place of residence.

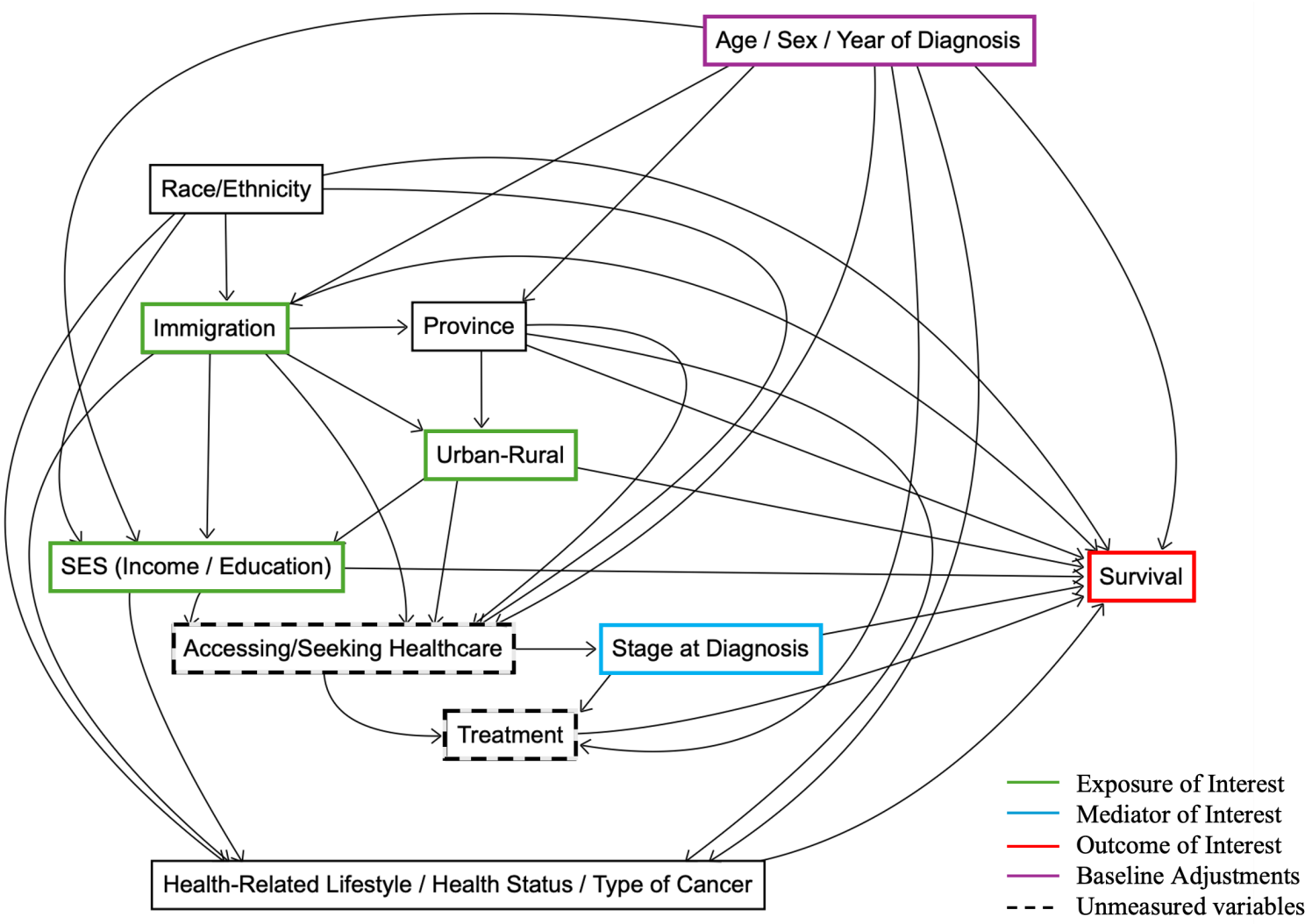


Figure 4.2.1. Directed acyclic graph (DAG)
Abbreviations: SES: socioeconomic status.

Statistical analyses

All statistical analyses were conducted using R software [32] and a population approach was adopted by incorporating survey weights and bootstrap weights for variance estimation. These weights were re-scaled using a 'pooled' approach to represent an average population [33] and normalized for integration into standard analytical software (Appendix D).[34, 35] Overall, the purpose of using survey weights was to account for the under-sampling of certain populations—such as immigrants—in the CCHS, who also represent the exposures of interest in our study. This led our inferences to be representative of individuals diagnosed with lung cancer between 2010 and 2016, among an average population residing in Canada between 2001 and 2017.

Frequencies, proportions, medians, and interquartile ranges were reported for patient characteristics overall and across exposure groups in the survey-weighted sample. Weighted Kaplan-Meier survival analyses were performed using R's *survival* and *survey* packages,[36, 37] incorporating transformed survey weights that were combined with stabilized inverse-probability-of-treatment weights (IPTW) (Appendix E). Exposure models used to generate IPTWs included confounding variables identified through our theoretical DAG. The number of categories for each variable was minimized to support model convergence and prevent extreme weights. The distribution of all weights was inspected for outliers, and the distribution of confounders was assessed before and after weighting.

We estimated IPTW- and survey-weighted survival curves for each exposure group and conducted weighted logrank tests. Differences in median survival times and percentile-based 95% confidence intervals (95% CI) based on combined bootstrap replicate weights were estimated. To protect CCHS respondents' confidentiality per Statistics Canada guidelines, we omitted cells with unweighted frequencies of less than 5 (or less than 15 for income-related analyses), rounded all frequencies to the nearest 5, and smoothed survival curves with penalized cubic regression splines.

We applied Lange et al.'s nested counterfactual approach for mediation analysis to examine the role of stage at diagnosis.[38] We first fit newly weighted and simulated datasets with McKeague and Sasieni's semi-parametric additive hazards models using the *timereg* package (Appendix F).[39, 40] From each natural effects model, we estimated the total effect (TE), which was decomposed into the pure natural direct effect (PNDE) and the total natural indirect effect (TNIE

– effect due to stage at diagnosis). The TNIE represents the expected change in hazards of death due to a change in cancer stage from its natural value under no exposure to that under exposure while remaining exposed. The percentage of TE mediated by advanced stage at diagnosis was also estimated. A 'full bootstrap' procedure was used to construct percentile-based 95% CIs.[41] For coherence, additive hazard models were fit using the original datasets to estimate hazard differences.

Two sensitivity analyses were conducted. First, we repeated our mediation analyses while excluding respondents from the 2015-2017 CCHS waves, which involved substantial methodological changes.[42] Second, we excluded individuals diagnosed with lung cancer before participating in the CCHS to mitigate any selection bias due to a survival advantage in these respondents. Additionally, we combined stage IIIa with IIIb and IV to examine the mediating role of locally advanced or metastatic cancers, since curative therapy rates drop from about 70% in stage II to approximately 40% in stage IIIa in Canada.[27, 43] Lastly, we conducted analyses stratified by sex due to potentially differing baseline hazards of death for men and women.

Results

Cohort characteristics

Of the 5,000 CCHS respondents (2000-2017 cycles) diagnosed with lung cancer in Canada between 2010 and 2016, we excluded those with missing information or stage 0 diagnosis, resulting in a sample of 4,430 (88.6%) respondents (Figure 4.3.1 in Appendix G).

In our survey-weighted sample, the median age at diagnosis was 71 years (IQR: 63-78); 47% of individuals were in the lowest income tercile, 42% had completed at least secondary education, 23% were immigrants, and 89% identified as White (Table 4.2.1). The median time from immigration to diagnosis was 44 years (IQR: 33-56). At diagnosis, 60% lived in large cities, and 52% had advanced-stage cancer. While 97% were ever-smokers, 69% reported excellent to good health in the CCHS, which was conducted within a median of 4.3 years (IQR: 1.6-7.8) from the time of diagnosis.

Stage at diagnosis varied slightly by income, education, and immigration status, but more by the urban-rural spectrum: 56% of residents in small cities were diagnosed at an advanced stage, compared to 52% in large cities, 46% in rural I areas, and 54% in rural II areas (Table 4.2.1). In

contrast, health status varied more by SES: 61.5% of low-income respondents reported excellent to good health, compared to 82% of high-income respondents.

Table 4.2.1. Characteristics of individuals in the survey-weighted sample

Characteristics	Overall sample (N=4,430)	Stratified sample, column percent										
		Family income tercile			Education		Immigrant status		Urban-rural spectrum			
	n (%)	1: Low (N=2,085)	2: Medium (N=1,520)	3: High (N=825)	<SS (N=1,525)	≥SS (N=2,905)	No (N=3,420)	Yes (N=995)	Large cities (N=2,675)	Small cities (N=775)	Rural I (N=655)	Rural II (N=330)
Socio-demographic, socio-economic, and geographic												
Age in years [A]												
<50	125 (2.9)	7.4 [K]	5.8 [K]	6.9 [K]	1.3	3.7	9.1	6.8	3.9	1.5	1.2	1.3
50-54	175 (3.9)				2.4	4.7	8.5	2.9	4.2	3.2	3.1	4.5
55-59	390 (8.9)	6.9	8.2	15.0	6.4	10.1	8.0	8.7	8.3	8.6	10.2	10.8
60-64	605 (13.7)	12.7	13.0	17.7	12.5	14.3	10.5	12.0	13.3	12.6	16.4	14.5
65-69	655 (14.8)	12.8	16.4	17.0	12.2	16.2	10.8	12.9	14.5	15.4	15.0	15.8
70-74	835 (18.9)	20.2	18.2	16.9	21.0	17.8	9.0	14.2	17.2	22.3	22.3	17.4
75-79	780 (17.6)	16.1	20.5	16.0	19.1	16.8	7.8	19.8	17.8	19.6	16.5	13.4
80-84	485 (10.9)	12.6	11.6	5.6	14.7	9.0	8.0	10.3	11.5	10.0	8.9	13.1
85-89	285 (6.4)	11.3 [K]	6.4 [K]	4.8 [K]	7.3	5.9	6.8	9.6	6.7	5.8	5.5	6.9
>=90	90 (2.0)				3.1	1.4	8.2	2.7	2.5	1.0	1.0	2.2
Median (interquartile range)	71 (63-78)	72 (64-79)	71 (64-78)	67 (61-75)	73 (65-79)	70 (62-77)	71 (63-78)	72 (62-79)	71 (63-78)	71 (64-78)	70 (63-77)	71 (63-78)
Female	2205 (49.8)	50.7	49.4	48.6	50.1	49.7	54.3	34.6	50.9	50.8	43.0	52.5
Marital status [B]												
Married	2410 (54.4)	40.6	63.4	72.7	47.5	58.0	51.7	63.8	51.1	61.3	59.6	54.4
Common-Law	160 (3.6)	3.9	2.5	5.0	3.2	3.9	3.9	2.7	3.5	2.8	4.1	5.7
Separated	110 (2.5)	3.1	1.1	3.4	3.0	2.2	2.9	1.0	2.3	2.8	3.2	2.1
Divorced	325 (7.3)	11.4	4.3	2.7	6.7	7.6	7.9	5.4	8.6	5.7	5.0	5.0
Widowed	850 (19.2)	25.3	15.7	10.1	26.3	15.4	20.0	16.3	20.1	17.4	17.1	20.1
Single	435 (9.8)	11.9	9.6	4.7	10.7	9.3	10.2	8.4	10.2	8.8	8.5	11.7
Missing	140 (3.2)	3.8	3.3	1.5	2.6	3.5	3.5	2.3	4.2	1.2	2.5	1.0
Family income tercile [C]												
1: Low	2085 (47.0)	100.0	--	--	63.9	38.2	45.9	50.8	46.9	44.0	48.3	52.7
2: Medium	1520 (34.4)	--	100.0	--	27.4	38.0	34.5	33.8	34.1	35.9	35.5	30.6
3: High	825 (18.6)	--	--	100.0	8.7	23.8	19.6	15.3	19.0	20.1	16.2	16.6
Education [D]												
Less than secondary school completion	1525 (34.4)	46.8	27.5	16.0	100.0	--	38.0	22.0	29.1	36.2	45.6	50.9
Secondary school completion and/or postsecondary education	2905 (65.6)	53.2	72.5	84.0	--	100.0	62.0	78.0	70.9	63.8	54.4	49.1

Table 4.2.1 (Continued). Characteristics of individuals in the survey-weighted sample

Characteristics	Overall sample (N=4,430)	Stratified sample, column percent										
		Family income tercile			Education		Immigrant status		Urban-rural spectrum			
	n (%)	1: Low (N=2,085)	2: Medium (N=1,520)	3: High (N=825)	<SS (N=1,525)	≥SS (N=2,905)	No (N=3,420)	Yes (N=995)	Large cities (N=2,675)	Small cities (N=775)	Rural I (N=655)	Rural II (N=330)
Immigrant status [D]												
Non-immigrant: Canadian-born	3430 (77.5)	75.7	77.8	81.5	85.6	73.2	100.0	--	69.6	84.9	92.3	94.8
Immigrant: foreign-born	995 (22.5)	24.3	22.2	18.5	14.4	26.8	--	100.0	30.4	15.1	7.7	5.2
Years since immigration at diagnosis [E]	44 (33-56)	38 (25-53)	47 (37-56)	47 (40-57)	48 (38-59)	42 (30-53)	--	44 (33-56)	40 (29-54)	50 (43-59)	49 (45-61)	52 (43-64)
Race/ethnicity												
White	3950 (89.2)	84.0	93.1	95.1	89.3	89.1	94.4	71.3	86.4	93.2	93.9	93.2
Non-White	460 (10.4)	15.3	6.7	4.7	10.0	10.6	5.2	28.2	13.1	6.5	5.9	6.7
Missing	20 (0.4)	0.6	0.2	0.2	0.7	0.3	0.3	0.6	0.5	0.3	0.2	0.1
Province [A, F]												
Alberta	510 (11.5)	14.3	8.6	9.9	9.9	12.3	12.3	8.8	12.1	9.7	9.4	14.5
British Columbia	825 (18.6)	19.4	19.3	15.2	17.1	19.4	16.3	26.3	16.9	28.1	13.8	19.3
Manitoba	240 (5.4)	5.4	5.6	5.0	6.8	4.7	6.2	2.6	5.4	2.5	7.3	8.7
New Brunswick	200 (4.6)	5.4	3.2	5.1	6.1	3.8	5.8	0.4	2.9	6.1	8.4	6.7
Nova Scotia	230 (5.2)	5.6 [L]	5.6 [L]	5.5 [L]	6.6	4.5	6.5	0.9	2.3	11.0	7.2	11.3
Prince Edward Island	30 (0.7)				0.8	0.6	0.8	0.2	0.0	2.1	1.9	0.2
Ontario	2055 (46.4)	41.9	48.9	53.2	42.5	48.4	42.5	59.9	56.2	35.1	37.9	10.6
Saskatchewan	175 (4.0)	4.2	4.2	3.1	4.6	3.7	5.0	0.5	3.1	2.9	3.8	14.5
Newfoundland and Labrador	150 (3.4)	3.2 [M]	4.6 [M]	3.1 [M]	5.1	2.5	4.3	0.3	1.1	1.9	10.3	12.0
YT, NWT, or NU	10 (0.3)				0.5	0.1	0.3	0.1	0.0	0.6	0.0	2.0
Urban-rural spectrum [A]												
Large cities: CMAs	2675 (60.4)	60.2	59.9	61.7	51.1	65.2	54.2	81.5	100.0	--	--	--
Small cities: CAs	775 (17.5)	16.3	18.3	18.8	18.4	17.0	19.1	11.7	--	100.0	--	--
Rural areas I: strong/moderate MIZs	655 (14.7)	15.1	15.2	12.8	19.5	12.2	17.5	5.1	--	--	100.0	--
Rural areas II: weak/no MIZs or territories	330 (7.4)	8.3	6.6	6.7	11.0	5.5	8.9	1.7	--	--	--	100.0
Cancer-related												
Year of lung cancer diagnosis												
2010	515 (11.6)	13.1	8.8	13.1	15.3	9.7	11.5	12.2	10.9	12.3	13.6	12.0
2011	420 (9.5)	9.0	9.8	10.4	9.6	9.5	10.2	7.3	8.8	10.7	9.4	13.0
2012	700 (15.8)	16.8	14.3	16.4	15.4	16.1	14.2	21.6	17.7	14.2	11.1	14.3
2013	585 (13.2)	12.4	13.8	14.2	13.0	13.3	13.5	12.1	12.4	13.0	15.8	15.2
2014	700 (15.8)	13.7	17.1	18.7	15.2	16.1	15.7	16.1	16.2	15.8	12.4	19.2
2015	730 (16.5)	16.5	18.8	12.1	14.8	17.4	16.3	17.2	17.5	16.0	15.8	11.1
2016	775 (17.5)	18.5	17.4	15.2	16.6	17.9	18.6	13.6	16.6	17.9	22.0	15.1

Table 4.2.1 (Continued). Characteristics of individuals in the survey-weighted sample

Characteristics	Overall sample (N=4,430)	Stratified sample, column percent										
		Family income tercile			Education		Immigrant status		Urban-rural spectrum			
	n (%)	1: Low (N=2,085)	2: Medium (N=1,520)	3: High (N=825)	<SS (N=1,525)	≥SS (N=2,905)	No (N=3,420)	Yes (N=995)	Large cities (N=2,675)	Small cities (N=775)	Rural I (N=655)	Rural II (N=330)
Lung cancer stage [A]												
Early stage: I-IIIa	2125 (48.0)	48.3	48.5	46.1	47.5	48.2	48.3	46.7	48.0	43.6	53.9	46.3
Stage I	1030 (23.3)	21.5	24.7	25.0	23.6	23.1	24.7	18.2	24.3	18.9	25.2	20.8
Stage II	460 (10.4)	11.8	9.9	7.8	7.9	11.7	8.6	16.7	11.1	7.4	10.8	11.4
Stage IIIa	630 (14.3)	15.0	23.8	13.2	16.0	13.4	15.0	11.8	12.6	17.3	17.9	14.1
Advanced stage: IIIb-IV	2305 (52.0)	51.7	51.5	53.9	52.5	51.8	51.7	53.3	52.0	56.4	46.1	53.7
Stage IIIb	320 (7.2)	6.5	6.9	9.4	5.8	7.9	7.0	7.8	7.4	7.6	5.3	8.3
Stage IV	1985 (44.8)	45.2	44.6	44.5	46.7	43.9	44.6	45.6	44.6	48.8	40.8	45.3
Lung cancer histology [A]												
Adenocarcinoma	2050 (46.4)	42.8	46.3	55.4	37.8	50.9	42.6	59.1	51.2	37.8	38.7	42.5
Squamous cell carcinoma	810 (18.3)	18.1	18.6	18.2	17.9	18.5	19.9	12.8	17.4	20.6	20.1	16.6
Large cell carcinoma	15 (0.3)	20.2 [N]	18.4 [N]	13.6 [N]	0.5	0.3	0.3	17.9 [N]	0.1	23.6 [N]	21.7 [N]	0.9
NSCLC, NOS	800 (18.0)				22.6	15.6	18.2		15.4			21.2
Small cell carcinoma	420 (9.4)	10.1	9.3	8.1	11.2	8.5	10.9	4.5	8.8	9.9	10.4	11.7
Unspecified	335 (7.5)	8.8	7.4	4.7	10.1	6.2	8.1	5.7	7.1	8.1	8.9	7.1
Health- and lifestyle-related												
Self-rated health [D]												
Excellent	415 (9.4)	8.9	8.0	13.4	7.7	10.3	8.5	12.5	10.1	9.1	9.0	5.4
Very good	1110 (25.1)	20.6	25.7	35.4	21.3	27.1	26.5	20.2	24.0	27.0	24.9	30.2
Good	1530 (34.5)	31.8	38.7	33.5	33.4	35.1	34.2	35.5	34.4	32.8	36.4	36.1
Fair	850 (19.2)	22.2	19.4	11.2	24.3	16.5	19.9	16.7	19.1	21.0	18.9	16.5
Poor	515 (11.6)	16.2	8.1	6.6	12.9	11.0	10.6	15.1	12.3	10.1	10.8	11.0
Missing	10 (0.2)	0.3	0.1	0.0	0.4	0.1	0.2	0.0	0.2	0.1	0.1	0.9
Has a regular medical doctor [D, G]												
Yes	4180 (94.4)	94.0	95.3	94.0	94.9	94.2	94.0	95.8	95.1	95.0	93.7	89.9
No	245 (5.5)	5.9	4.6	5.9	5.0	5.8	5.9	4.1	4.9	5.0	6.3	10.1
Missing	5 (0.1)	0.1	0.0	0.1	0.1	0.0	0.1	0.0	0.1	0.1	0.0	0.0
Visited a doctor in the past year [D, H]												
Yes	3500 (79.1)	78.4	81.6	75.9	76.9	80.2	78.3	81.6	79.2	78.5	81.5	74.5
No	500 (11.2)	11.1	9.4	14.9	13.2	10.2	12.0	8.7	11.0	11.1	10.0	16.1
Missing	430 (9.7)	10.5	8.9	9.2	9.9	9.6	9.7	9.7	9.8	10.4	8.6	9.4
Had a flu shot [D, I]												
Ever	3005 (67.9)	64.3	70.7	71.7	65.5	69.2	69.4	62.7	68.3	71.8	66.6	58.0
Never	1130 (25.5)	28.1	22.7	24.3	26.0	25.3	24.7	28.3	25.1	21.2	27.9	34.4
Missing	290 (6.6)	7.6	6.5	4.1	8.6	5.5	5.9	9.1	6.6	7.0	5.4	7.5

Table 4.2.1 (Continued). Characteristics of individuals in the survey-weighted sample

Characteristics	Overall sample (N=4,430)	Stratified sample, column percent										
		Family income tercile			Education		Immigrant status		Urban-rural spectrum			
	n (%)	1: Low (N=2,085)	2: Medium (N=1,520)	3: High (N=825)	<SS (N=1,525)	≥SS (N=2,905)	No (N=3,420)	Yes (N=995)	Large cities (N=2,675)	Small cities (N=775)	Rural I (N=655)	Rural II (N=330)
Alcohol drinking in the past year [D]												
Yes	3075 (69.4)	61.7	73.0	82.4	60.5	74.1	70.5	65.6	69.9	66.9	72.0	66.6
No	1325 (29.9)	37.9	25.6	17.3	39.2	25.0	29.1	32.4	29.2	32.7	27.6	33.1
Missing	30 (0.7)	0.4	1.4	0.3	0.3	0.9	0.3	2.1	0.9	0.4	0.4	0.4
Alcohol binging in the past year [D]												
Yes	2045 (46.2)	42.6	47.8	52.4	39.7	49.6	44.7	51.3	47.8	43.4	46.9	38.3
No	1330 (30.1)	38.3	25.7	17.5	39.5	25.2	29.4	32.6	29.4	32.9	28.0	33.1
Missing	1050 (23.7)	19.2	26.5	30.1	20.9	25.2	25.9	16.1	22.8	23.7	25.1	28.6
Smoking status [D]												
Never smoker	305 (6.9)	8.2	6.2	5.0	6.4	7.1	4.4	15.4	7.9	6.2	4.7	4.7
Former smoker	2115 (47.8)	40.3	51.1	60.5	44.3	49.6	46.9	50.9	48.0	48.9	46.7	45.6
Current smoker	2005 (45.2)	51.5	42.7	34.1	49.2	43.1	48.6	33.7	44.1	44.8	48.6	49.3
Missing	5 (0.1)	0.0	--	0.4	0.0	0.1	0.1	0.0	0.1	0.0	0.0	0.4
Smoking intensity [D]												
Pack-years for former smoker [E]	30 (8.5-50)	31 (8.0-53.8)	29.3 (10.8-50.8)	28.0 (7.5-45.6)	33.0 (14.7-57.0)	27.2 (6.5-47.0)	32.0 (14.2-51.6)	12.0 (0.0-43.5)	26.0 (6.0-46.3)	33.5 (16.4-53.8)	37.5 (16.5-57.8)	33.2 (14.1-51.2)
Pack-years for current smokers [E]	38 (24.0-54.0)	38.3 (24.0-55.0)	38.8 (22.5-53.0)	34.5 (24.7-51.3)	43.0 (26.6-58.7)	35.3 (22.9-51.3)	40.0 (25.0-55.0)	30.6 (14.2-42.6)	36.4 (23.5-54.0)	37.5 (22.5-52.5)	43.9 (27.3-56.3)	36.9 (22.5-51.3)
Missing	275 (6.2)	5.9	7.2	5.0	6.1	6.3	6.6	4.8	6.2	6.8	4.1	9.0
Time-related												
Year of survey												
2000	400 (9.1)	9.3	9.5	7.6	9.0	9.1	9.2	8.7	9.2	8.1	9.7	8.9
2003	385 (8.7)	8.2	9.6	8.6	9.2	8.5	8.9	8.2	9.0	9.5	6.8	8.4
2005	370 (8.4)	7.7	8.6	9.7	9.0	8.0	9.0	6.2	8.0	9.2	7.8	10.4
2007	440 (9.9)	9.4	10.8	9.7	8.8	10.5	10.0	9.8	9.7	9.7	11.0	10.1
2008	470 (10.6)	10.8	8.7	13.6	10.1	10.8	10.1	12.3	12.0	8.8	8.4	8.0
2009	400 (9.0)	8.4	10.9	6.9	9.0	9.0	8.9	9.6	7.6	12.0	7.6	16.1
2010	400 (9.0)	10.4	6.9	9.4	11.3	7.8	9.6	7.0	9.5	8.8	8.2	7.6
2011	360 (8.1)	8.8	7.4	7.9	8.3	8.0	7.9	8.8	8.3	8.6	6.6	8.0
2012	340 (7.7)	7.9	7.3	7.9	9.4	6.8	7.7	7.8	7.0	7.4	11.7	6.6
2013	320 (7.3)	7.2	6.5	8.8	5.5	8.2	7.7	6.0	7.0	6.3	10.4	5.8
2014	285 (6.5)	5.7	9.3	3.1	4.7	7.4	4.9	11.8	7.4	5.1	5.3	4.2
2015 or 2016 [J]	175 (3.9)	6.1 [O]	4.4 [O]	6.7 [O]	3.9	3.9	4.2	2.8	3.6	4.4	4.0	5.1
2017 [J]	75 (1.7)				1.8	1.7	1.9	1.1	1.5	2.2	2.5	0.8

Table 4.2.1 (Continued). Characteristics of individuals in the survey-weighted sample

Characteristics	Overall sample (N=4,430)	Stratified sample, column percent										
		Family income tercile			Education		Immigrant status		Urban-rural spectrum			
	n (%)	1: Low (N=2,085)	2: Medium (N=1,520)	3: High (N=825)	<SS (N=1,525)	≥SS (N=2,905)	No (N=3,420)	Yes (N=995)	Large cities (N=2,675)	Small cities (N=775)	Rural I (N=655)	Rural II (N=330)
Years from survey to lung cancer diagnosis [E]	4.3 (1.6-7.8)	4.0 (1.6-7.8)	4.6 (1.8-8.0)	4.0 (1.6-7.7)	4.0 (1.7-7.8)	4.4 (1.5-7.9)	32.0 (14.2-51.6)	12.0 (0.0-43.5)	4.4 (1.6-8.0)	3.9 (1.6-7.7)	4.1 (1.4-7.7)	4.5 (1.7-7.2)
Months of follow-up after lung cancer diagnosis [E]	11.9 (3.2-29.7)	10.9 (3.0-27.5)	11.6 (3.0-29.7)	16.5 (4.5-36.9)	9.2 (2.7-27.9)	12.8 (3.7-30.2)	40.0 (25.0-55.0)	30.6 (14.2-42.6)	11.7 (3.2-29.7)	10.2 (2.6-25.3)	13.4 (3.9-31.7)	11.6 (3.6-32.4)
Died during follow-up	3230 (73.0)	75.7	71.2	69.4	75.9	71.5	71.8	76.9	73.0	76.8	67.9	73.9

[A] Measured at the time of lung cancer diagnosis.
[B] Measured in the year prior to the year of lung cancer diagnosis.
[C] Population-level terciles of equivalized after-tax family income at 2 years prior to lung cancer diagnosis, that are specific by province, urban-rural residence (CMA/CA/Rural/Territories), and calendar year.
[D] Measured at the time of CCHS survey.
[E] Median (interquartile range)
[F] Quebec not included because it did not contribute cancer incidence data to the CCR after 2010
[G] Includes any health professional for respondents of CCHS in 2015+
[H] Accounts for family doctor, pediatrician, or general practitioner. This question was optional for provinces in CCHS cycles 2011 and 2017.
[I] Excludes the H1N1 vaccine.
[J] For CCHS cycles 2015+, only files that combined 2 annual cycles included respondents from the territories. The file for CCHS cycle 2017 did not include respondents from territories.

Combined categories due to vetting rules related to cells with n<5 (n<15 for income-related analyses):
[K] <50 with 50-54 and 85-59 with >=90.
[L] Nova Scotia with Prince Edward Island.
[M] Newfoundland and Labrador with Yukon, Northwest Territories or Nunavut
[N] Large cell carcinoma with NSCLC, NOS
[O] CCHS waves 2015 and 2016 with 2017.

Abbreviations: SS: secondary school; CCHS: Canadian Community Health Survey; CCR: Canadian Cancer Registry; CMAs: Census Metropolitan Areas; CAs: Census Agglomerations; MIZs: Metropolitan Influence Zones; YT: Yukon; NWT: Northwest Territories; NU: Nunavut.

Overall survival

With a median follow-up of 11.9 months during our study period, we observed 3,230 deaths and a median overall survival of 11.9 months (95% CI: 10.3 to 13.5) (Table 4.2.1).

Survival curves by income showed no clear differences (logrank p-value: 0.240); however, all curves crossed, violating the proportional hazards assumption required for the logrank test. In contrast, potential differences in survival curves across education were observed (logrank p-value: 0.0789) (Figure 4.2.2). Median survival differences between low- and high-income respondents, and those without versus with secondary education, were -4.3 months (95% CI: -9.3 to 1.7) and -3.5 months (95% CI: -5.0 to -0.2), respectively (Table 4.2.2). The corresponding hazard differences (TEs) were 44.8 deaths (95% CI: -3.3 to 101.2) and 39.6 deaths (95% CI: -7.4 to 79.0) per 10,000 person-months. Little evidence for differences by immigration status was found, with a logrank p-value of 0.347 (Figure 4.2.2) for survival curves, no median survival difference (-0.13 months, 95% CI: -4.1 to 3.0), and the highest uncertainty in hazard difference (33.0 deaths per 10,000 person-months, 95% CI: -31.8 to 109.0). Differences in survival curves were observed across the urban-rural spectrum (logrank p-value: 0.0215) (Figure 4.2.2). Small compared to large city residents had almost no median survival difference (-0.82 months, 95% CI: -4.4 to 1.6), but the largest hazard difference (50.1 additional deaths per 10,000 person-months, 95% CI: -3.7 to 107.2). Hazard differences for rural I and II areas showed greater uncertainty in the direction of the estimates with values of -27.5 deaths (95% CI: -70.4 to 29.0) and -33.2 deaths (95% CI: -95.9 to 28.9) per 10,000 person-months.

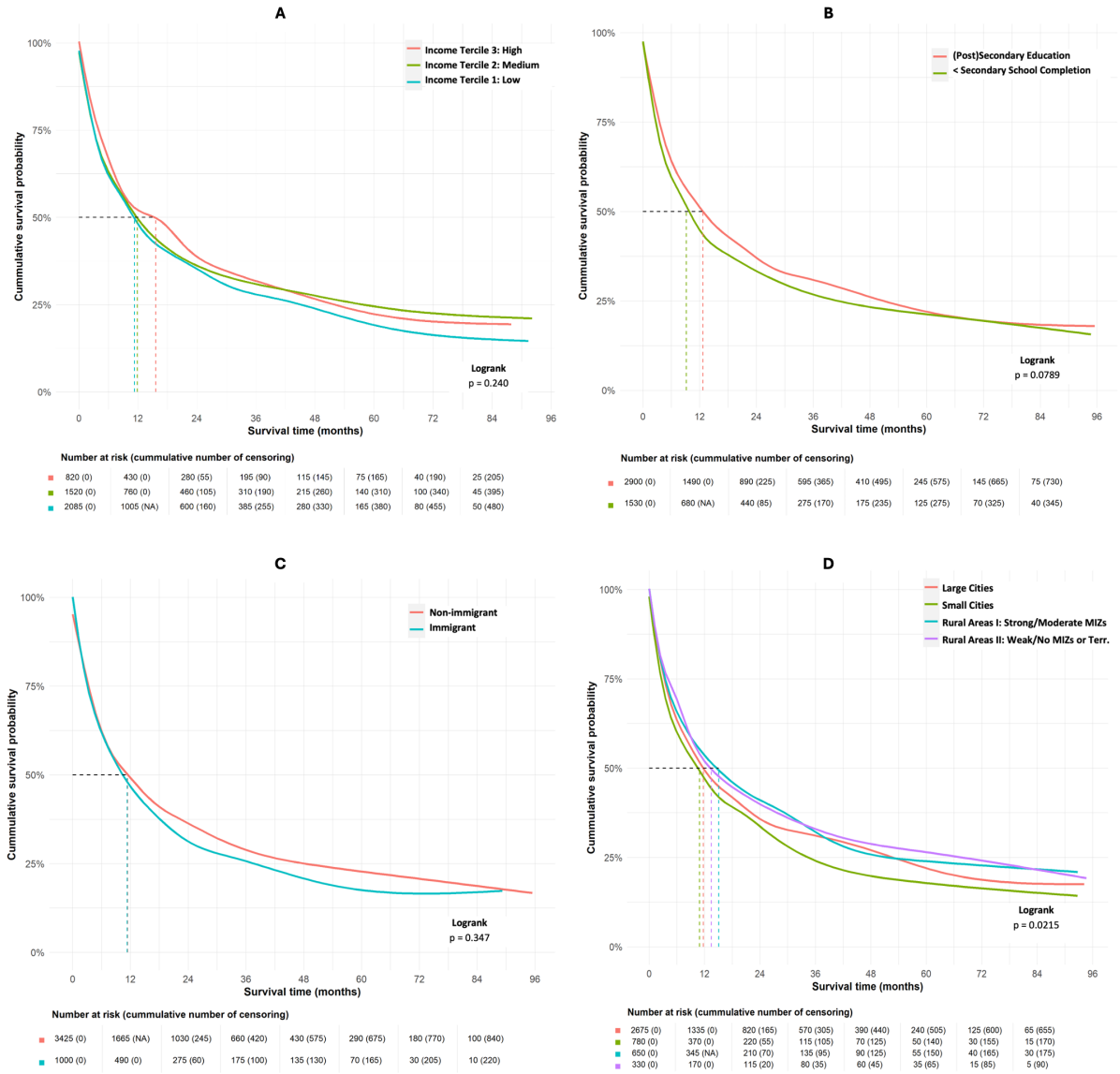


Figure 4.2.2. Kaplan-Meier survival curves for income (A), education (B), immigration status (C), and urban-rural spectrum (D).
Survey-weighted N = 4430.

The dashed lines represent the median survival time. References groups are in red.
All curves accounted for survey weights and inverse-probability treatment weights. IPTW models accounted for age (4 categories: <65, 65-74, 75-79, and ≥80), sex, and year of diagnosis (2 categories: 2010-2013, and 2014-2016). In addition, IPTW models for income and education as exposures accounted for province (4 categories: 1- British Columbia; 2- Manitoba and Alberta; 3- New Brunswick, Newfoundland and Labrador, and Ontario; and 4- Saskatchewan, Nova Scotia, Prince Edward Island, and Territories) and rurality (3 categories: large cities, small cities, and rural areas or territories). The model for urban-rural spectrum of residence as exposure accounted for province, and the model for immigration as exposure accounted for race/ethnicity (White and non-White, missing were imputed as White [n=20]). Provinces grouped according to previously reported 1-year survival rates.[13]

Abbreviations: MIZ: Metropolitan Influence Zones.

Table 4.2.2. Survival anlayses incorporating survey and IPTW weights

Exposure variables	Median survival months (95% CI)	Median survival months Difference (95% CI)	Hazard difference, per 10,000 person-months (95% CI)
Family income terciles [A, B]			
1: Low	11.33 (8.84, 12.78)	-4.30 (-9.32, 1.74)	44.84 (-3.30, 101.21)
2: Medium	11.89 (9.59, 13.98)	-3.75 (-8.07, 2.71)	15.77 (-31.35, 69.81)
3: High	15.64 (9.72, 19.58)	REF	REF
Education [A]			
Less than secondary school completion	9.20 (8.28, 11.95)	-3.52 (-5.04, -0.21)	39.55 (-7.43, 79.02)
Secondary school completion and/or postsecondary education	12.71 (10.86, 13.93)	REF	REF
Immigration status [A]			
Non-immigrant (Canadian-born)	11.47 (9.77, 12.78)	REF	REF
Immigrant (foreign-born)	11.33 (7.95, 13.63)	-0.13 (-4.12, 3.01)	32.96 (-31.76, 108.99)
Urban-rural spectrum [A]			
Large cities (CMAs)	11.76 (9.35, 13.62)	REF	REF
Small cities (CAs)	10.94 (7.92, 12.62)	-0.82 (-4.44, 1.63)	50.10 (-3.66, 107.21)
Rural areas I (strong/moderate MIZs)	15.08 (10.72, 18.89)	3.32 (-1.37, 7.26)	-27.51 (-70.38, 29.00)
Rural areas II (weak/no MIZs or territories)	13.50 (9.10, 20.24)	1.74 (-2.79, 9.37)	-33.25 (-95.92, 28.87)

[A] Survey-weighted N = 4430. IPTW models accounted for age (4 categories: <65, 65-74, 75-79, and ≥80), sex, and year of diagnosis (2 categories: 2010-2013, and 2014-2016). In addition, IPTW models for income and education as exposures accounted for province (4 categories: 1- British Columbia; 2- Manitoba and Alberta; 3- New Brunswick, Newfoundland and Labrador, and Ontario; and 4- Saskatchewan, Nova Scotia, Prince Edward Island, and Territories) and rurality (3 categories: large cities, small cities, and rural areas or territories). The model for urban-rural spectrum of residence as exposure accounted for province, and the model for immigration as exposure accounted for race/ethnicity (White, and non-White, missing were imputed as White [n=20]). Provinces grouped according to previously reported 1-year survival rates.[13]

[B] Population-level terciles of equivalized after-tax family income at 2 years prior to lung cancer diagnosis, that are specific by province, urban-rural residence (CMA/CA/Rural/Territories), and calendar year.

Abbreviations: CMAs: Census Metropolitan Areas; CAs: Census Agglomerations; MIZs: Metropolitan Influence Zones.

Mediation analysis

Small versus large cities had one of the largest TNIE estimates with a relatively narrow confidence interval: 20.0 deaths (95% CI: -11.1 to 47.5) per 10,000 person-months (red bars in Figure 4.2.3 and Table 4.3.5 in Appendix H). The TE for small versus large cities was also the largest and consistent in direction with the TNIE, and point estimates suggest that 40.7% of the effect of living in small vs. large cities is mediated by advanced stage at diagnosis. We observed small TNIEs for high vs. low income or education: -11.5 deaths (95% CI: -42.7 to 26.5) and 8.3 deaths (95% CI: -22.9 to 35.4) per 10,000 person-months, respectively.

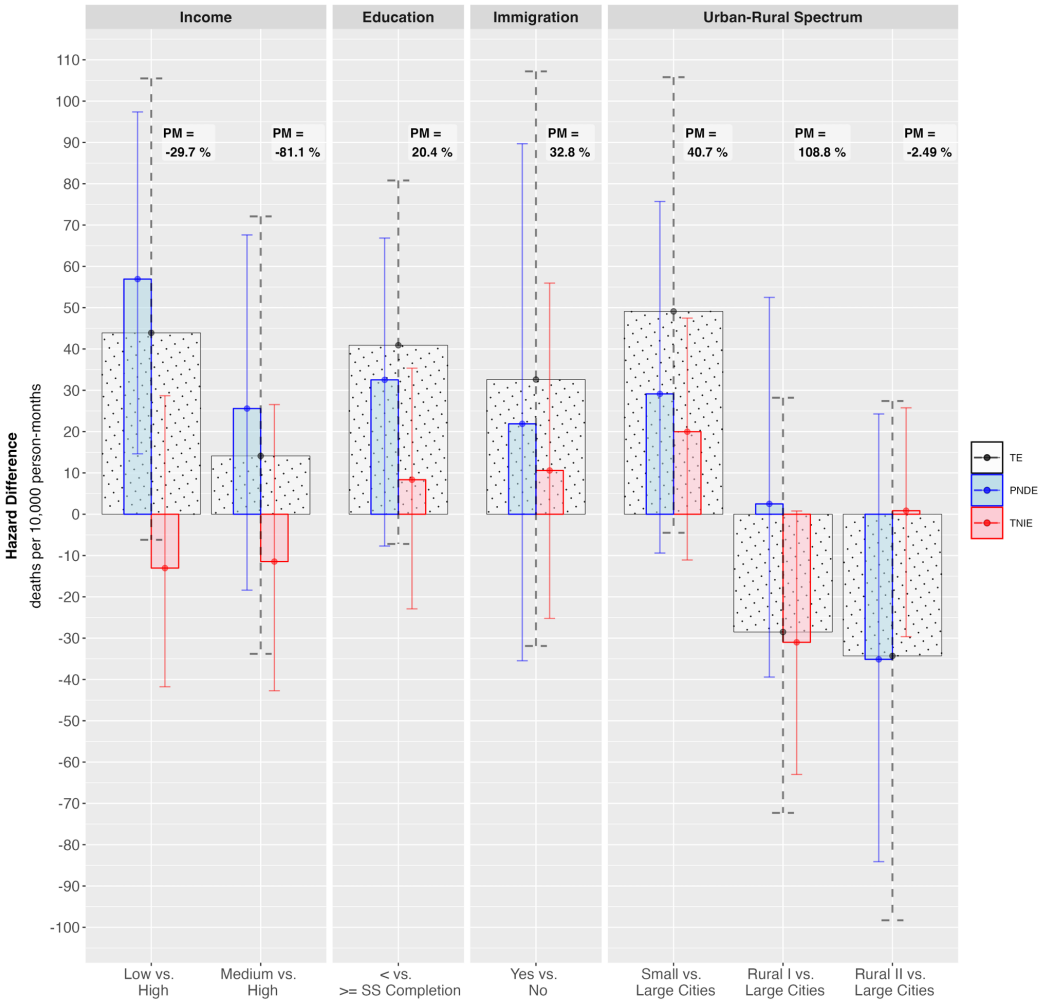


Figure 4.2.3. Main mediation analyses for income, education, immigration status, and urban-rural spectrum, including survey, IPTW, and mediation weights.
Survey-weighted N = 4430.

The grey dotted bars represent total effects, the blue bars represent the pure natural direct effects, the red bars represent the total natural indirect effects, and the corresponding whiskers represent the 95% confidence intervals.

The total effect (TE) is the sum of the pure natural direct effect (PNDE) and the total natural indirect effect (TNIE). PNDE quantifies the average hazard difference due to change in exposure (e.g., switching from high to low income), while stage at diagnosis (i.e., the mediator) assumes the value it would have naturally under no exposure (e.g., high income). TNIE quantifies the average hazard difference from a change in the mediator from the natural value it would have taken under no exposure to the natural value under exposure, while remaining exposed.

IPTW models used in mediation analyses accounted for age (4 categories: <65, 65-74, 75-79, and ≥80), sex, and year of diagnosis (2 categories: 2010-2013, and 2014-2016). In addition, IPTW models for income and education as exposures accounted for province (4 categories: 1- British Columbia; 2- Manitoba and Alberta; 3- New Brunswick, Newfoundland and Labrador, and Ontario; and 4- Saskatchewan, Nova Scotia, Prince Edward Island, and Territories) and rurality (3 categories: large cities, small cities, and rural areas or territories). The model for urban-rural spectrum of residence as exposure accounted for province, and the model for immigration as exposure accounted for race/ethnicity (White and non-White, missing were imputed as White [n=20]). Provinces grouped according to previously reported 1-year survival rates.[13]

Abbreviations: SS: secondary school; TE: total effects; PNDE: pure natural direct effects; TNIE: total natural indirect effects; PM: percent mediated.

Secondary analyses

When the mediator was redefined, the TNIE for small vs. large cities increased from 20.0 deaths (95% CI: -11.1 to 47.5) to 46.5 deaths (95% CI: 12.7 to 74.7) per 10,000 person-months (Figure 4.2.4 and Table 4.3.6 in Appendix H). Sex-stratified analyses indicate that survival inequalities based on income are larger among females, while city size differences are larger among males (Figure 4.2.4 and Table 4.3.6). For example, the difference between small and large city residents was 77.3 deaths (95% CI: -13.5 to 176.3) and 32.1 deaths (95% CI: -26.8 to 85.5) per 10,000 person-months for males and females, respectively. TNIEs for the urban-rural spectrum remained consistent across sex.

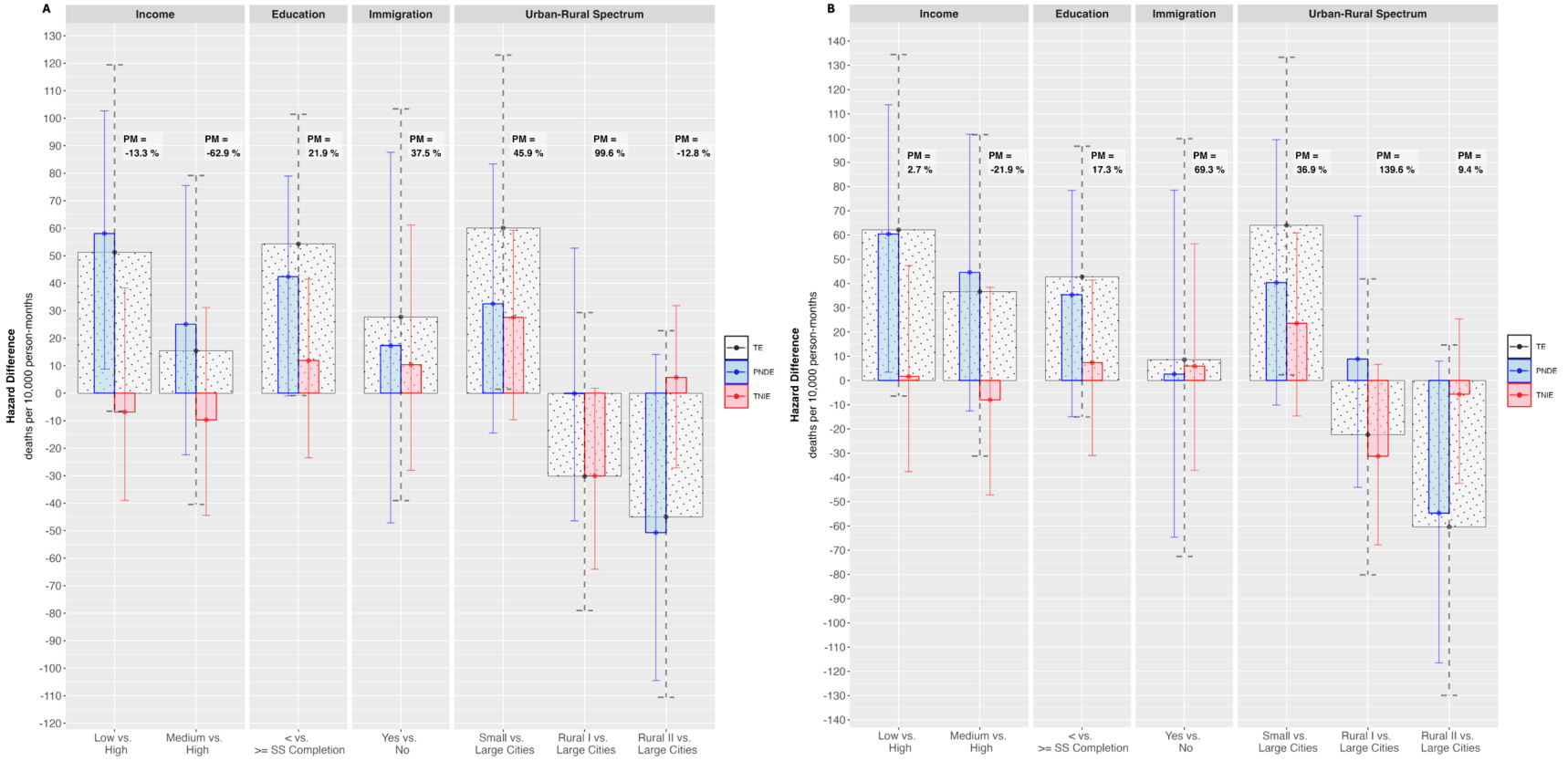


Figure 4.2.4. Secondary mediation analyses for income, education, immigration status, and urban-rural spectrum, including survey, IPTW, and mediation weights.

Panel A: CCHS respondents from waves 2000 to 2014; survey-weighted N = 4140.

Panel B: CCHS respondents from waves 2000 to 2016 diagnosed with lung cancer after CCHS participation; survey-weighted N = 4010.

See footnotes as in Figure 4.2.3.

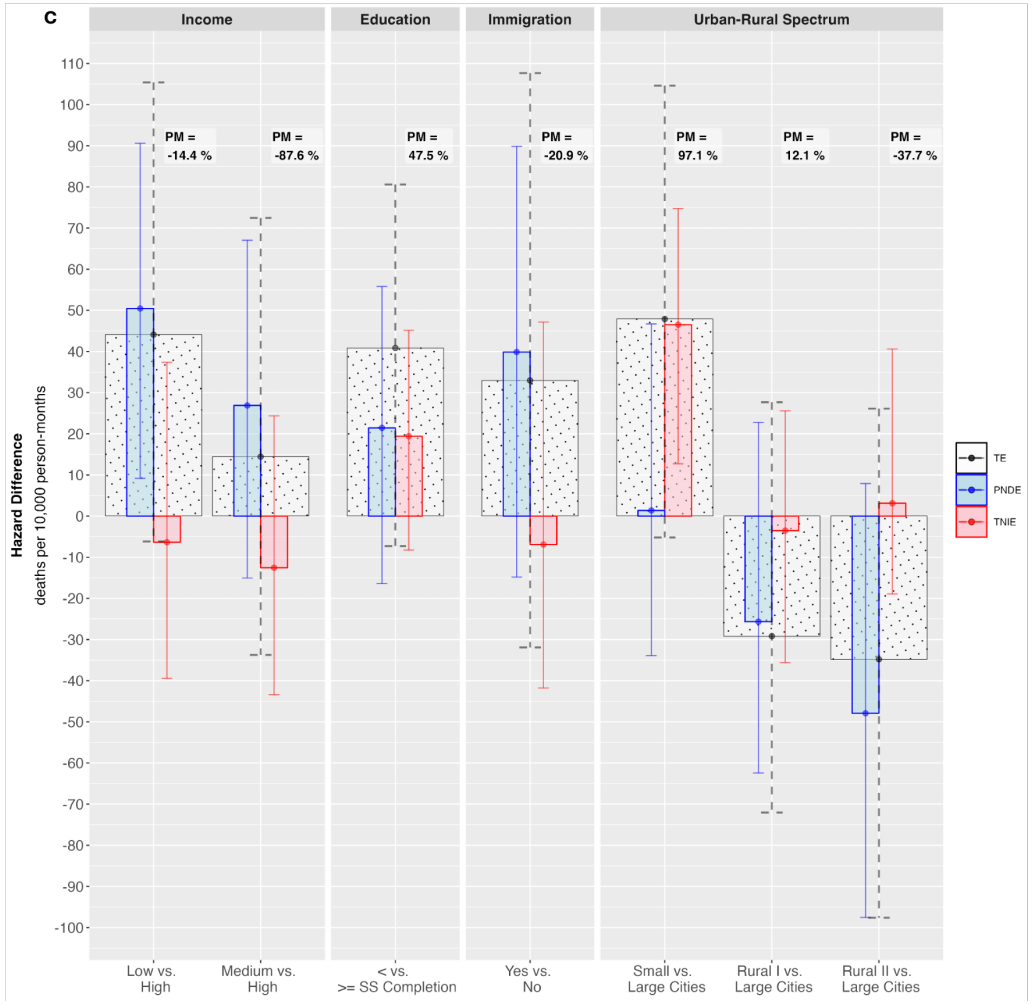


Figure 4.2.4 (continued). Secondary mediation analyses for income, education, immigration status, and urban-rural spectrum, including survey, IPTW, and mediation weights.

Panel C: CCHS respondents from waves 2000 to 2017; advanced stage (mediator) redefined as stages IIIa, IIIb, and IV; survey-weighted N = 4430.
See footnotes as in Figure 4.2.3.

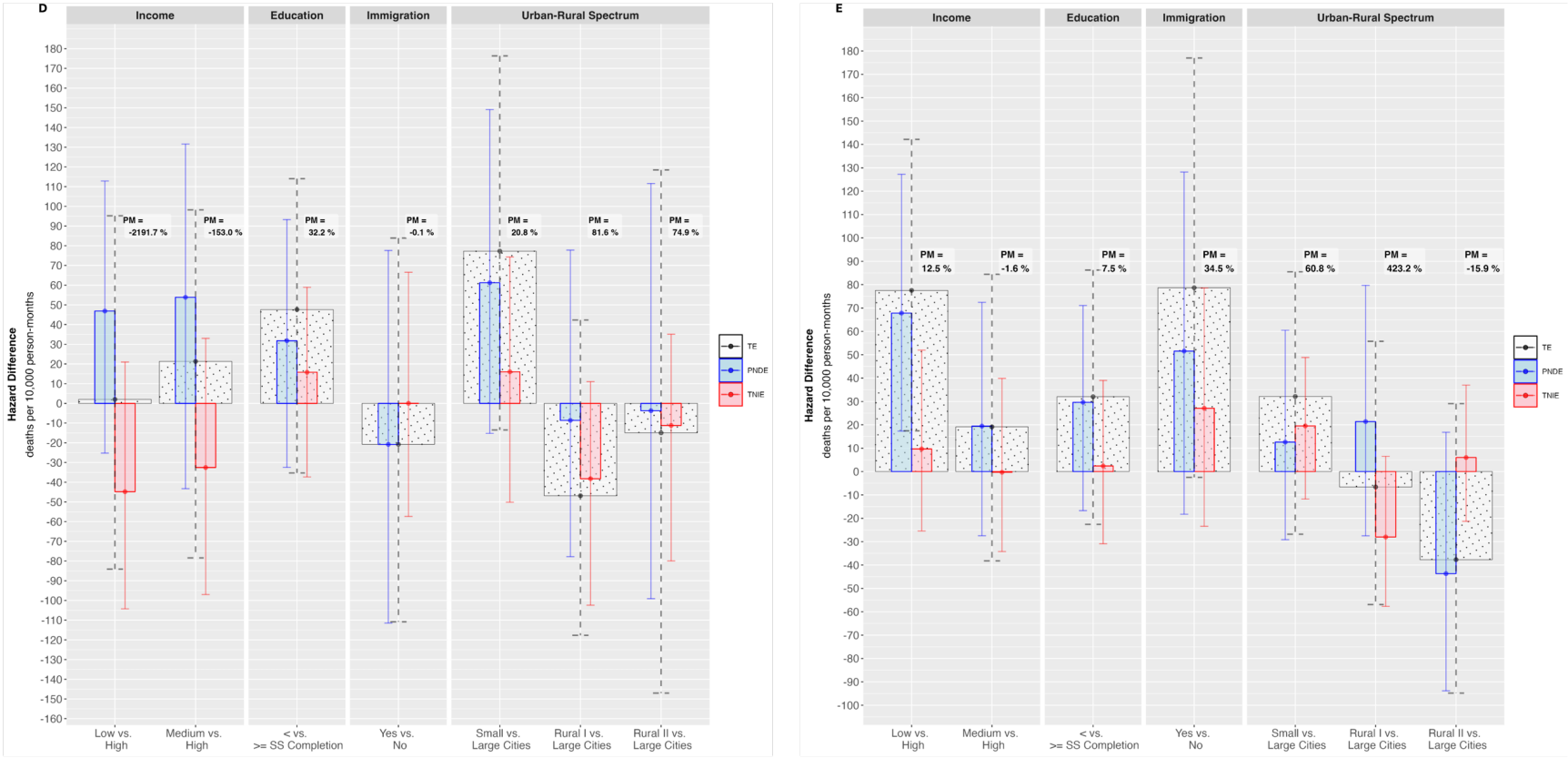


Figure 4.2.4 (continued). Secondary mediation analyses for income, education, immigration status, and urban-rural spectrum, including survey, IPTW, and mediation weights.

Panel D: CCHS respondents from waves 2000 to 2017; males; survey-weighted N = 1955.

Panel E: CCHS respondents from waves 2000 to 2017; females; survey-weighted N = 2475.

See footnotes as in Figure 4.2.3.

Discussion

We observed meaningful survival inequalities by individual-level SES and by place of residence on an urban-rural spectrum among lung cancer patients in Canada. Contrary to SES-based inequalities, a large portion of the survival inequality in small compared to large urban cities was mediated by stage at diagnosis.

We estimated median survival among respondents with the lowest versus highest levels of income and education to be 4.3 and 3.5 months shorter, respectively. Our estimates of differences in survival outcomes qualitatively align with those from previous studies on inequalities in lung cancer survival based on education and both individual- and area-level income in Canada.[2, 5-11, 13, 44, 45] We observed the smallest differences in the distribution of advanced stage at diagnosis across income and education, and our small TNIE estimates suggest that stage at diagnosis may have a limited role in mediating SES-based survival inequalities. Studies on area-based SES and family income in Canada found no association with lung cancer stage at diagnosis,[9, 11, 46] including no mediation of SES-based survival inequalities by stage at diagnosis.[9] It has been suggested that treatment and comorbidities may explain SES-based inequalities in lung cancer survival.[20, 21]

In our second sensitivity analysis, we observed a difference of only 8.6 deaths per 10,000 person-months by immigrant status, providing little evidence of inequalities and suggesting that the estimated hazard differences in our main analysis were impacted by selection bias. Although immigrants, who constitute 22% of Canada's population,[47] often underutilize cancer prevention and care services,[48-53] immigrants of 10 years or less generally exhibit longer cancer survival than non-immigrants, a benefit that diminishes with time since immigration.[54, 55] In our study, 75% of immigrants had been in Canada for at least 33 years. One recent study in Ontario reported longer survival among immigrants compared to non-immigrants, though 75% of the immigrants in the study had been residing in Canada for less than 24 years.[56] Another national-level study, using census respondents who developed lung cancer, also found a survival advantage for immigrants over non-immigrants; however, no adjustments were made for differences in sex.[45] It has been shown that there are more immigrant women than men among seniors in Canada,

including within the lung cancer population—a trend consistent with the general Canadian senior population.[56, 57]

Compared to large city residents in our study, those in rural areas I and II had 27.5 and 33.2 fewer deaths per 10,000 person-months, while small city residents had 50.1 more deaths per 10,000 person-months. Canadian studies on various cancers, including two on lung cancer,[2, 8] showed mixed results for dichotomous rural vs. urban survival: worse outcomes,[58-61] better outcomes,[8, 62, 63] or no difference.[2, 64, 65] In general, individuals living in remote or rural areas are part of underserved communities that face barriers to accessing cancer care;[1, 66] however, cancer-related outcomes can vary by cancer type. For lung cancer, studies have reported no significant differences in the distribution of advanced-stage diagnoses or receipt of curative surgery between urban and rural areas [12]—two of the strongest determinants of lung cancer survival. Our survival estimates for rural areas are consistent with these previous findings, although they vary with the degree of remoteness and are associated with the greatest uncertainty in our study.

Of seven studies on non-lung cancers using granular urban-rural definitions, four found worse survival outcomes in small vs. large urban areas,[58, 62, 63, 67] while others found better outcomes [65] or no significant differences.[59, 64] One of the studies reported that patient case mix, which included stage at diagnosis, partly explained poorer survival in a small urban region compared to a larger one.[67, 68] Finally, one study in Ontario demonstrated significantly lower survival among individuals diagnosed with lung cancer residing in small urban areas relative to those in large urban areas despite adjusting for stage at diagnosis.[69] The TNIEs we observed for small vs. large cities suggest that advanced stage at diagnosis mediates a large portion of survival inequalities, and this remained consistent in our secondary analyses. In a post hoc analysis, redefining our mediator and excluding respondents who were diagnosed with lung cancer prior to responding to the CCHS survey, we found a TE of 63.3 deaths per 10,000 person-months (95% CI: 2.1 to 132.7) and a TNIE of 48.4 deaths (95% CI: 12.1 to 87.4), indicating up to 76.5% of the total effect of city size is mediated by diagnosis of locally advanced or metastatic cancers (Figure 4.3.2 and Table 4.3.7 in Appendix H). This aligns with our observation of similar median survival times between individuals in small and large urban regions, where over 50% are diagnosed at an advanced stage. If advanced stage at diagnosis is the primary driver of survival inequalities, as our

findings suggest, then minimal differences in median survival would be expected—given the uniformly poor prognosis associated with late-stage lung cancer.

Canada's urban landscape is polarized, with large cities experiencing only growth, while 21% of smaller cities faced shrinkage from 2006 to 2011.[70] Shrinking cities face employment and social program cutbacks, along with aging populations that require increased healthcare.[70] Age, health status, stage at diagnosis, and access to treatment may contribute to survival differences in small vs. large urban areas.[58, 62] However, quality of care and imbalances between healthcare demand and resources could also factor in and remain largely unexplored.[58, 62] Residents of small urban and rural areas, relative to those in large urban areas, experience more difficulties in accessing specialized physicians and cancer experts despite having higher cancer prevalence.[71-73] Moreover, residents from small urban areas report higher unmet healthcare needs than those in rural areas.[71] Limited resources may not meet healthcare demands in small urban areas, but may suffice in rural areas with smaller populations. Rural areas also have more family physicians providing cancer care compared to small urban areas [74] and are known to have shorter wait times for seeing specialists and receiving cancer diagnoses compared to urban areas.[75, 76]

This study has limitations. While semi-parametric additive hazard models assume constant hazard differences over time, goodness-of-fit tests and cumulative regression coefficient plots for the urban-rural spectrum variable in our study suggest time-invariant direct effects and attenuating indirect effects (Table 4.3.8 in Appendix I). Indirect effects diminished around 750 days and were largest in the first 150 days post-diagnosis (Tables 4.3.8 and 4.3.9 in Appendix I), aligning with the median survival time for stage IV lung cancer patients of 4.8 months.[28] Our main TNIE estimates are likely underestimated for early survival times (Supplementary Table 4.3.10 in Appendix I).

Our survival estimates rely on no unmeasured exposure-outcome confounding. Our natural effect estimates additionally rely on no unmeasured exposure-mediator or mediator-outcome confounding, and no exposure-induced confounding. Due to limited cell counts, we accounted for province instead of race/ethnicity or immigration status in some analyses, potentially leading to residual confounding. Estimating natural effects for upstream exposures like immigration status and place of residence is prone to exposure-induced confounding (Figure 4.2.1). This is less

concerning for immigration status since it lacked an association with survival, and for place of residence, for which downstream SES-based inequalities were not mediated by stage at diagnosis. However, estimates obtained from our additive hazard models had considerable uncertainty, especially the TNIE estimates. If stage at diagnosis truly mediates inequalities by income and education, alternative mediation methods may be needed to address exposure-induced confounding when investigating upstream exposures.

This study has significant strengths. It is one of the first to comprehensively examine survival outcomes by individual-level socioeconomic and sociodemographic characteristics of lung cancer patients in Canada. We are also one of the first to apply counterfactual-based methods to time-to-event data to elucidate the role of stage at diagnosis in lung cancer survival inequalities. Previous studies have inferred the role of stage at diagnosis by comparing stage distribution across exposures, associations of exposures with survival stratified by stage, or survival associations with and without stage adjustments, often focusing on survival at one point in time, which may be more biased or uninformative.[21] Our methods, which account for complex survey design and data linkage combined with sensitivity analyses, provide more reliable population-level estimates.

Conclusion

Urban areas are not homogeneous, and incorporating urban-rural gradients should be a priority in future research on area-based inequalities in lung cancer outcomes. Interventions that effectively target advanced stage at diagnosis may have a greater impact on reducing survival inequalities by place of residence than by SES.

Funding

This research was funded by a doctoral research award (Frederick Banting and Charles Best Canada Graduate Scholarships) granted to S.Q. by the Canadian Institutes of Health Research (CIHR funding number FBD-175860). S.Q is also a recipient of the CRDCN Emerging Scholars Grant supporting the undertaking of this research in an RDC. E.S. holds McGill University's William Dawson Scholar award, which supports her research.

Conflict of interest

None declared

Ethics approval

Ethics approval was granted by McGill University's Faculty of Medicine and Health Sciences Institutional Review Board on April 12, 2022 (study number: A04-M30-22A).

Author contributions

Samia Qureshi designed the study and analysis plan, managed the data, conducted the analyses, interpreted the results, and drafted the manuscript. **Mireille E Schnitzer**, **Arijit Nandi**, and **Erin Strumpf** contributed to the interpretation of the results, reviewed the manuscript, and approved the final draft. Erin Strumpf supervised the study.

Data availability

Data used in this study were provided by Statistics Canada and accessed through one or more of the RDCs (Research Data Centres) in the Canadian Research Data Centre Network (CRDCN). All analyses were conducted with data from the following data-linkage project in which Canadian Population Health Survey data (CCHS Annual) was integrated with multiple datasets: '*Linkage of the Canadian Community Health Survey to mortality, cancer, hospital administrative files, Census (short-form), and tax data*'. Because of the confidential nature of these microdata, they cannot be shared. Researchers in Canada working at one of CRDCN's member institutions can access the data at no additional cost to the researcher. Other researchers will have to pay cost-recovery to access the data. Access to the data is subject to a background check and research approval process. The protocols for data access, including fees for researchers at non-CRDCN institutions, can be found on the CRDCN website.

Acknowledgements

The analysis presented in this paper was conducted at the Quebec Interuniversity Centre for Social Statistics (QICSS), which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the QICSS are made possible by the financial or in-kind support of the Social Sciences and Humanities Research Council (SSHRC), the Canadian Institutes of Health Research (CIHR), the Canada Foundation for Innovation (CFI), Statistics Canada, the Fonds de recherche du Québec - Société et culture (FRQSC), the Fonds de recherche du Québec - Santé (FRQS), and the Quebec universities. The views expressed in this paper are those of the authors, and not necessarily those of the CRDCN or its partners.

References

(1) Canadian Cancer Society's Advisory Committee in collaboration with the Canadian Cancer Society, Statistics Canada and the Public Health Agency of Canada. Canadian Cancer Statistics 2023. Toronto, ON: Canadian Cancer Society; 2023.

(2) Habbous S, Khan Y, Langer DL, 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. DOI: https://doi.org/10.4103/atm.atm_283_20

(3) Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer survival: a review. Ann Oncol. 2006;17(1):5-19. DOI: https://doi.org/10.1093/annonc/mdj007

(4) Finke I, Behrens G, Weisser L, Brenner H, Jansen L. Socioeconomic Differences and Lung Cancer Survival-Systematic Review and Meta-Analysis. Front Oncol. 2018;8:536-. DOI: https://doi.org/10.3389/fonc.2018.00536

(5) Mackillop WJ, Zhang-Salomons J, Groome PA, Paszat L, Holowaty E. Socioeconomic status and cancer survival in Ontario. J Clin Oncol. 1997;15(4):1680-9. DOI: https://doi.org/10.1200/JCO.1997.15.4.1680

(6) 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. 1999;17(7):2244-. DOI: https://doi.org/10.1200/JCO.1999.17.7.2244

(7) Zhang–Salomons J, Qian H, Holowaty E, Mackillop WJ. Associations Between Socioeconomic Status and Cancer Survival: Choice of SES Indicator May Affect Results. Ann Epidemiol. 2006;16(7):521-8. DOI: https://doi.org/10.1016/j.annepidem.2005.10.002

(8) Groome PA, Schulze KM, Keller S, Mackillop WJ. Demographic Differences between Cancer Survivors and those who Die Quickly of their Disease. Clin Oncol. 2008;20(8):647-56. DOI: https://doi.org/10.1016/j.clon.2008.05.006

(9) Booth CM, Li G, Zhang-Salomons J, Mackillop WJ. The impact of socioeconomic status on stage of cancer at diagnosis and survival. Cancer. 2010;116(17):4160-7. DOI: https://doi.org/10.1002/cncr.25427

(10) Shah M, Parmar A, Chan KKW. Socioeconomic disparity trends in diagnostic imaging, treatments, and survival for non-small cell lung cancer 2007-2016. Cancer medicine. 2020;9(10):3407-16. DOI: https://doi.org/10.1002/cam4.2978

(11) Ruan Y, Heer E, Warkentin MT, et al. The association between neighborhood-level income and cancer stage at diagnosis and survival in Alberta. Cancer. 2023. DOI: https://doi.org/10.1002/cncr.35098

(12) Canadian Partnership Against Cancer. Lung cancer and equity: A focus on income and geography. Toronto, ON.: Canadian Partnership Against Cancer; 2020.

(13) Canadian Cancer Society's Advisory Committee. Canadian Cancer Statistics: A 2020 special report on lung cancer. Toronto, ON: Canadian Cancer Society; 2020.

(14. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. WHO Document Production Services; 2010.

(15) Glymour MM, Avendano M, Kawachi I. 17Socioeconomic Status and Health. In: Berkman LF, Kawachi I, Glymour MM, editors. Social Epidemiology: Oxford University Press; 2014. p. 0.

(16) Vang Z, Sigouin J, Flenon A, Gagnon A. The healthy immigrant effect in Canada: A systematic review. Population Change and Lifecourse Strategic Knowledge Cluster Discussion Paper Series/Un Réseau stratégique de connaissances Changements de population et parcours de vie Document de travail. 2015;3(1):4. DOI:

(17) Krieger N. Defining and investigating social disparities in cancer: critical issues. Cancer Causes Control. 2005;16(1):5-14. DOI: 10.1007/s10552-004-1251-5

(18) Fayet Y, Praud D, Fervers B, et al. Beyond the map: evidencing the spatial dimension of health inequalities. International Journal of Health Geographics. 2020;19(1):46. DOI: 10.1186/s12942-020-00242-0

(19) Davis LE, Mahar AL, Strumpf EC. Agreement between individual and neighborhood income measures in patients with colorectal cancer in Canada. J Natl Cancer Inst. 2023;115(5):514-22. DOI: https://doi.org/10.1093/jnci/djad017

(20) Forrest LF, Sowden S, Rubin G, White M, Adams J. Socio-economic inequalities in stage at diagnosis, and in time intervals on the lung cancer pathway from first symptom to

treatment: systematic review and meta-analysis. Thorax. 2017;72(5):430. DOI: https://doi.org/10.1136/thoraxjnl-2016-209013

(21) Afshar N, English DR, Milne RL. Factors Explaining Socio-Economic Inequalities in Cancer Survival: A Systematic Review. Cancer Control. 2021;28:10732748211011956. DOI: https://doi.org/10.1177/10732748211011956

(22) Jiang Z, VanderWeele TJ. When Is the Difference Method Conservative for Assessing Mediation? Am J Epidemiol. 2015;182(2):105-8. DOI: https://doi.org/10.1093/aje/kwv059

(23) Afshar N, English DR, Milne RL. Rural–urban residence and cancer survival in high-income countries: A systematic review. Cancer. 2019;125(13):2172-84. DOI: https://doi.org/10.1002/cncr.32073

(24) Lin RS, Plevritis SK. Comparing the benefits of screening for breast cancer and lung cancer using a novel natural history model. Cancer Causes Control. 2012;23(1):175-85. DOI: https://doi.org/10.1007/s10552-011-9866-9

(25) Postal Code OM Conversion File Plus (PCCF+) [Internet]. Government of Canada. 2023 [cited 2023-12-06]. Available from: https://www150.statcan.gc.ca/n1/en/catalogue/82F0086X#wb-auto-2.

(26) 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;17(6):1471-4. DOI: https://doi.org/10.1245/s10434-010-0985-4

(27) Vinod SK, Wai E, Alexander C, Tyldesley S, Murray N. Stage III non–small-cell lung cancer: population-based patterns of treatment in British Columbia, Canada. J Thorac Oncol. 2012;7(7):1155-63. DOI:

(28) Seung SJ, Hurry M, Walton RN, Evans WK. Real-world treatment patterns and survival in stage IV non-small-cell lung cancer in Canada. Curr Oncol. 2020;27(4):e361-e7. DOI: https://doi.org/10.3747/co.27.6049

(29) Eberhardt WEE, De Ruysscher D, Weder W, et al. 2nd ESMO Consensus Conference in Lung Cancer: locally advanced stage III non-small-cell lung cancer. Ann Oncol. 2015;26(8):1573-88. DOI: https://doi.org/10.1093/annonc/mdv187

(30) Ellison LF, Saint-Jacques N. Five-year cancer survival by stage at diagnosis in Canada. Health Rep. 2023;34(1):3-15. DOI: https://doi.org/10.25318/82-003-x202300100001-eng

(31) Hajizadeh M, Johnston GM, Manos D. Socio-economic inequalities in lung cancer incidence in Canada, 1992–2010: results from the Canadian Cancer Registry. Public Health. 2020;185:189-95. DOI: https://doi.org/10.1016/j.puhe.2020.04.023

(32) R Core Team. R: A Language and Environment for Statistical Computing. 2023. Vienna, Austria. Available from: https://www.R-project.org/.

(33) Thomas S, Wannell B. Combining cycles of the Canadian community health survey. Health Rep. 2009;20(1):53. DOI: https://www150.statcan.gc.ca/n1/en/pub/82-003-x/2009001/article/10795-eng.pdf?st=Ak1RqveT

(34) Rao J, Wu C, Yue K. Some recent work on resampling methods for complex surveys. Survey methodology. 1992;18(2):209-17. DOI: https://www150.statcan.gc.ca/n1/en/pub/12-001-x/1992002/article/14486-eng.pdf?st=YEBRFJO4

(35) Gagné C, Roberts G, Keown L-A. Weighted estimation and bootstrap variance estimation for analyzing survey data: How to implement in selected software. The Research Data Centres Information and Technical Bulletin. 2014;6(1):1-72. DOI: https://www150.statcan.gc.ca/n1/pub/12-002-x/2014001/article/11901-eng.htm

(36) Therneau T. A Package for Survival Analysis in R. 2021. Available from: https://CRAN.R-project.org/package=survival.

(37) Lumley T. survey: analysis of complex survey samples. 2023.

(38) Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. Am J Epidemiol. 2012;176(3):190-5. DOI: https://doi.org/10.1093/aje/kwr525

(39) McKeague IW, Sasieni PD. A partly parametric additive risk model. Biometrika. 1994;81(3):501-14. DOI: https://doi.org/10.2307/2337122

(40) Scheike TH, Zhang M-J. Analyzing competing risk data using the R timereg package. Journal of statistical software. 2011;38(2). DOI: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3375021/

(41) Reaume M, McIsaac MA. Natural direct and indirect effects for survival outcomes in a complex survey setting. Kingston, Canada: Queen's University; 2017.

(42) Canadian Community Health Survey - Annual Component (CCHS): Summary of changes. [Internet]. Government of Canada. 2022. Available from: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getMainChange&Id=3359.

(43) Moore S, Leung B, Wu J, Ho C. Population-based analysis of curative therapies in stage II non-small cell lung cancer: the role of radiotherapy in medically inoperable patients. Radiat Oncol. 2020;15(1):23. DOI: https://doi.org/10.1186/s13014-020-1466-y

(44) Stavraky KM, Skillings JR, Stitt LW, Gwadry-Sridhar F. The effect of socioeconomic status on the long-term outcome of cancer. J Clin Epidemiol. 1996;49(10):1155-60. DOI: https://doi.org/10.1016/0895-4356(96)00179-5

(45) Malagón T, Botting-Provost S, Moore A, El-Zein M, Franco EL. Inequalities in relative cancer survival by race, immigration status, income, and education for 22 cancer sites in

Canada, a cohort study. medRxiv. 2024:2024.05.31.24307976. DOI: 10.1101/2024.05.31.24307976

(46) Lofters AK, Gatov E, Lu H, et al. Lung Cancer Inequalities in Stage of Diagnosis in Ontario, Canada. Current Oncology. 2021;28(3):1946-56. DOI: https://doi.org/10.3390%2Fcurroncol28030181

(47) Statistics Canada. 2016 Census of Population: Immigration and ethnocultural diversity. In: Canada Go, editor.: Minister of IIndustry; 2017.

(48) Maxwell CJ, Kozak JF, Desjardins-Denault SD, Parboosingh J. Factors important in promoting mammography screening among Canadian women. Canadian Journal of Public Health. 1997;88:346-50. DOI: https://doi.org/10.1007%2FBF03403903

(49) Maxwell CJ, Bancej CM, Snider J. Predictors of mammography use among Canadian women aged 50–69: findings from the 1996/97 National Population Health Survey. CMAJ. 2001;164(3):329-34. DOI: https://www.cmaj.ca/content/164/3/329

(50) Datta GD, Blair A, Sylvestre M-P, Gauvin L, Drouin M, Mayrand M-H. Cervical cancer screening in Montreal: Building evidence to support primary care and policy interventions. Prev Med. 2018;111:265-71. DOI: https://doi.org/10.1016/j.ypmed.2018.02.037

(51) Bacal V, Blinder H, Momoli F, Wu KY, McFaul S. Is Immigrant Status Associated With Cervical Cancer Screening Among Women in Canada? Results From a Cross-Sectional Study. Journal of Obstetrics and Gynaecology Canada. 2019;41(6):824-31.e1. DOI: https://doi.org/10.1016/j.jogc.2018.07.010

(52) Blair A, Gauvin L, Schnitzer ME, Datta GD. The Role of Access to a Regular Primary Care Physician in Mediating Immigration-Based Disparities in Colorectal Screening: Application of Multiple Mediation Methods. Cancer Epidemiol Biomarkers Prev. 2019;28(4):650-8. DOI: https://doi.org/10.1158/1055-9965.Epi-18-0825

(53) Canadian Partnership Against Cancer. Examining disparities in cancer control: A system performance special focus report. Canadian Partnership Against Cancer; 2014.

(54) Stevenson JKR, Cheung MC, Earle CC, et al. Chinese and South Asian ethnicity, immigration status, and clinical cancer outcomes in the Ontario Cancer System. Cancer. 2018;124(7):1473-82. DOI: 10.1002/cncr.31231

(55) Cheung MC, Earle CC, Fischer HD, et al. Impact of Immigration Status on Cancer Outcomes in Ontario, Canada. J Oncol Pract. 2017;13(7):e602-e12. DOI: https://doi.org/10.1200/JOP.2016.019497

(56) Ruco A, Lofters AK, Lu H, et al. Lung cancer survival by immigrant status: a population-based retrospective cohort study in Ontario, Canada. BMC Cancer. 2024;24(1):1114. DOI: 10.1186/s12885-024-12804-7

(57) Table 98-10-0347-01 Immigrant status and period of immigration by gender and age: Canada, provinces and territories. [Internet]. 2023.

(58) Bosma NA, Tilley D, Batra A, Cheung WY. Characterizing Urban-Rural Differences in Colon Cancer Outcomes: A Population-based Analysis Based on Travel Distance to Cancer Center. Am J Clin Oncol. 2020;43(7). DOI: https://doi.org/10.1097/coc.0000000000000703

(59) Gorey KM, Fung KY, Luginaah IN, et al. Cancer Survival in Ontario, 1986–2003. Canadian Journal of Public Health. 2008;99(1):12-6. DOI: https://doi.org/10.1007/BF03403733

(60) Raju RS, Coburn N, Liu N, et al. A population-based study of the epidemiology of pancreatic cancer: a brief report. Curr Oncol. 2015;22(6):e478-84. DOI: 10.3747/co.22.2653

(61) Hallet J, Law CHL, Karanicolas PJ, Saskin R, Liu N, Singh S. Rural-urban disparities in incidence and outcomes of neuroendocrine tumors: A population-based analysis of 6271 cases. Cancer. 2015;121(13):2214-21. DOI: https://doi.org/10.1002/cncr.29338

(62) Lee B, Goktepe O, Hay K, et al. Effect of place of residence and treatment on survival outcomes in patients with diffuse large B-cell lymphoma in British Columbia. Oncologist. 2014;19(3):283-90. DOI: https://doi.org/10.1634/theoncologist.2013-0343

(63) Olson RA, Nichol A, Caron NR, et al. Effect of Community Population Size on Breast Cancer Screening, Stage Distribution, Treatment Use and Outcomes. Canadian Journal of Public Health. 2012;103(1):46-52. DOI: https://doi.org/10.1007/BF03404068

(64) Kim JD, Firouzbakht A, Ruan JY, et al. Urban and rural differences in outcomes of head and neck cancer. The Laryngoscope. 2018;128(4):852-8. DOI: https://doi.org/10.1002/lary.26836

(65) Loree JM, Javaheri KR, Lefresne SV, et al. Impact of Travel Distance and Urban-Rural Status on the Multidisciplinary Management of Rectal Cancer. The Journal of Rural Health. 2017;33(4):393-401. DOI: https://doi.org/10.1111/jrh.12219

(66) Canadian Partnership Against Cancer. 2019-2029: Canadian Strategy for Cancer Control. Doing together what cannot be done alone.; 2019.

(67) Webber C, Brundage M, Hanna TP, et al. Explaining regional variations in colon cancer survival in Ontario, Canada: a population-based retrospective cohort study. BMJ Open. 2022;12(9):e059597. DOI: https://doi.org/10.1136/bmjopen-2021-059597

(68) Lucas G, Bielska IA, Fong R, Johnson AP. Rural-urban differences in use of health care resources among patients with ankle sprains in Ontario. Canadian Journal of Rural Medicine. 2018;23(1):7+. DOI: https://go.gale.com/ps/i.do?id=GALE%7CA521592168&sid=googleScholar&v=2.1&it=r

&linkaccess=abs&issn=12037796&p=HRCA&sw=w&userGroupName=anon%7E47ce4c3d&aty=open-web-entry

(69) Mullin ML, Shellenberger J, Whitehead M, et al. Characterizing Regional Variability in Lung Cancer Outcomes across Ontario-A Population-Based Analysis. Curr Oncol. 2022;29(12):9640-59. DOI: 10.3390/curroncol29120757

(70) Hartt M, Hollander J. City size and academic focus: Exploring trends in Canadian urban geography, planning and policy literature. Canadian Journal of Urban Research. 2018;27(1):1-13. DOI: https://cjur.uwinnipeg.ca/index.php/cjur/article/view/111

(71) Sibley LM, Weiner JP. An evaluation of access to health care services along the rural-urban continuum in Canada. BMC Health Serv Res. 2011;11(1):1-11. DOI: https://doi.org/10.1186/1472-6963-11-20

(72) Rai Y, Zheng S, Chappell H, Pulandiran M, Jones J. Kidney cancer survivorship care: Patient experiences in a Canadian setting. Can Urol Assoc J. 2020;14(11):E560-e7. DOI: https://doi.org/10.5489/cuaj.6217

(73) Terashima M, Rainham DG, Levy AR. A small-area analysis of inequalities in chronic disease prevalence across urban and non-urban communities in the Province of Nova Scotia, Canada, 2007-2011. BMJ Open. 2014;4(5):e004459. DOI: https://doi.org/10.1136/bmjopen-2013-004459

(74) Pong RW, Pitblado JR, Canadian Institute for Health I. Geographic distribution of physicians in Canada : beyond how many and where. Ottawa, Ontario: Canadian Institute for Health Information 2005. Report No.: 1553927389.

(75) Harrington DW, Wilson K, Rosenberg M, Bell S. Access granted! barriers endure: determinants of difficulties accessing specialist care when required in Ontario, Canada. BMC Health Serv Res. 2013;13(1):146. DOI: https://doi.org/10.1186/1472-6963-13-146

(76) Porter G, Urquhart R, Kendell C, Bu J, McConnell Y, Grunfeld E. Timely access and quality of care in colorectal cancer: a population-based cohort study using administrative data. BMC Res Notes. 2013;6(1):355. DOI: https://doi.org/10.1186/1756-0500-6-355

4.3. Supplementary materials

4.3.1. Appendix A

Databases used and information retrieved

Table 4.3.1. Data sources

	Years	Information extracted	Timing of information relative to lung cancer diagnosis
CCHS	Waves 2000/01, 2003/04, 2005/06, and 2007 to 2017	- education - immigration - race/ethnicity - health status - health-related behaviors: seeking healthcare alcohol drinking smoking - time between survey and lung cancer diagnosis	- before diagnosis for respondents from waves 2000/01 to 2009 - before or after diagnosis for respondents from waves 2010 to 2016 - after diagnosis for respondents from wave 2017
CCR incidence TMF	2010-2016	- age - sex - date of lung cancer diagnosis - cancer stage - tumour histology - postal code	- at the time of diagnosis
T1FF [A]	2005-2015	- income - marital status	- income: 2 years before diagnosis - marital status: 1 year before diagnosis
CVSD	2010-2017	- time of death	- after or at the time of diagnosis
PCCF+ [B]	Census (SAS program versions): - census 2011 (6d) - census 2016 (7e)	- province - Statistical Area Classification type	- at the time of diagnosis

[A] Information from T1FF files corresponding to the year of lung cancer diagnosis was excluded due to incomplete records, likely due to respondents dying within the same year.

[B] PCCF+ version 6d was used for postal codes in diagnosis years 2010 to 2014 and version 7e for postal codes in diagnosis years 2015-2016.

Abbreviations: CCHS: Canadian Community Health Survey; CCR incidence TMF: Canadian Cancer Registry incidence Tabulation Master File; T1FF: T1 Family File; CVSD: Canadian Vital Statistics Database; PCCF+: Postal Code Conversion File Plus

1. The Canadian Community Health Survey (CCHS): Waves 2000/01, 2003/04, 2005/06, and 2007 to 2017

The CCHS is an annual (biennial until 2007) survey that collects self-reported information on demographics, health status, healthcare utilization, and health behaviours of the Canadian population aged 12 years and older (excluding individuals in the Canadian Forces, special care facilities, prisons, First Nations reserves or Crown lands, northern health regions of Quebec, and those without a place of residence). We have harmonized CCHS variables of interest across all survey cycles to account for changes or renaming of survey questions, especially due to the redesign of CCHS in 2015, which introduced a new sampling methodology and content. Smoking information was harmonized with a modified version of the method included in the *cashflow* package in R.[1] Since 2013, the definition of alcohol binging for female respondents has slightly differed, now requiring the consumption of four drinks on one occasion instead of five or more drinks. CCHS data includes individual-level weights and 500-1000 bootstrap weights that account for survey design and non-response. The survey weights are used to obtain population-level point estimates, while bootstrap survey weights are applied for calculating variances related to those estimates.

The timing of information obtained from the CCHS concerning the lung cancer diagnosis varied: information was collected before diagnosis for respondents from the 2000/01 to 2009 waves, after diagnosis for those from the 2017 wave, or either before or after diagnosis for those from the 2010 to 2017 waves 2016.

2. The Canadian Cancer Registry Incidence Tabulation Master File (CCR Incidence TMF): 2010-2016

The CCR is a Statistics Canada database that consolidates data collected and reported by each of the Canadian provincial and territorial cancer registries, standardizing information on all cancers diagnosed in Canada since 1992. Staging data has been systematically compiled in the CCR from 2010 and onward. However, Quebec's cancer incidence data was not submitted to the CCR during this period, specifically in the files that were part of the linkage project used in this study. Our study focused on individuals who responded to the CCHS surveys from 2000 to 2017 and were linked to a primary lung cancer diagnosis in the CCR between 2010 and 2016.

Selection of lung cancer tumours

For all tumours reported in 2010 and onwards, multiple primary cancer diagnoses were identified in the CCR Incidence TMF files according to the Surveillance, Epidemiology, and End Results (SEER) rules.[2] We selected all tumours in the CCR Incidence TMF with SEER grouping code 0302 for lung and bronchus cancers. Based on the International Classification of Diseases for Oncology, third edition (ICD-O-3), the SEER grouping for lung and bronchus cancers includes all tumours reported with topography code C34 and histology codes 8000 to 9049, 9056 to 9139, and 9141 to 9589. Histology codes for mesotheliomas (9050–9055), Kaposi sarcomas (9140), and hematopoietic and lymphoid neoplasms (9590–9992) are categorized in other SEER groups; therefore, these cancers are excluded from our analyses.

Handling of duplicates

The CCR Incidence TMF includes multiple primary tumours, as individuals can be diagnosed with more than one primary cancer in their lifetime. Among all CCHS respondents identified with lung cancer in the CCR Incidence TMF, approximately 0.8% had multiple primary lung cancer tumours, with a maximum of two tumours per respondent. We retained only one tumour entry per respondent based on the following criteria: 1) when the primary lung cancer tumours were identical based on the date of diagnosis and cancer stage, either one of the two observations was retained, 2) when the diagnosis dates were identical, but the stages of diagnosis differed, the observation with the most advanced cancer stage, or the one that does not have missing information on cancer stage was retained, 3) when the diagnosis dates were different and 30 days apart, or when the diagnosis dates were within 30 days, and the stages of diagnoses were identical, the observation with the earliest date of diagnosis was retained, 4) when the diagnosis dates were within 30 days, but the stages of diagnoses differed, the observation with the most advanced cancer stage or the observation that is not missing information on cancer stage was retained.

Defining lung cancer histologic type

Similar to the Canadian Cancer Society's Advisory Committee's histological groupings, we used the ICD-O-3 histological codes to categorize lung cancer tumours into seven types of histology: 1) squamous cell carcinoma, 2) adenocarcinoma, 3) small cell carcinoma, 4) large cell carcinoma, 5) non-small cell lung cancer, not otherwise specified, 6) sarcomas and other specified neoplasms, and 7) unspecified.[3]

The Canadian Cancer Society's Advisory Committee's histological groupings	
Histologic type	**ICD-O-3 Histology code**
Squamous cell carcinoma	8050-8078
Adenocarcinoma	8140, 8211, 8230, 8231, 8250-8260, 8323, 8480-8490, 8550-8552, 8570-8574, 8576
Small cell carcinoma	8040-8045
Large cell carcinoma	8011-8012, 8014-8031, 8035, 8310
Non-small cell lung cancer, not otherwise specified (NOS)	Excluding codes above: 8010-8576
Sarcomas and other specified neoplasms	8800-8811, 8830, 8840-8921, 8990-8991, 9040-9044, 9120-9133, 9150, 9540-9581, 8815, 8980
Unspecified	8000-8005

3. The Annual Income Estimates for Census Families and Individuals file (T1 Family File or T1FF): 2005-2015

The T1FF is compiled by the Canada Revenue Agency and includes information on sources of income for individuals identified in personal income tax returns or the Canada Child Benefit program (i.e., tax filers and non-filers). T1FF files corresponding to the year of lung cancer diagnosis were not used due to incomplete records, likely caused by respondents dying in the same year.

4. The Canadian Vital Statistics – Death Database (CVSD): 2010-2017

The CVSD provides information on all deaths in Canada by consolidating information collected and reported by all provincial and territorial vital statistics registries. It includes information such as demographics, date of death, and cause of death.

5. The Postal Code Conversion Files – Plus (PCCF+)

The PCCF+ is a SAS program that links single 6-digit postal codes in Canada to standard geographic areas as defined by Statistics Canada.[4] These standard geographic areas are used to produce census data and some statistics that are also found in the PCCF+ programs (e.g., neighbourhood-level socioeconomic measures). Hence, there are different versions of the PCCF+

programs to account for changes in standard geographic areas and statistics by census year. Postal codes can be associated with more than one geographic area, and the PCCF program (not PCCF+) links each postal code to a single geographic area where most dwellings are located. In contrast, the PCCF+ program uses census-derived weights for population-weighted random allocation (i.e., probabilistic assignment) of postal codes to a geographic area. The PCCF+ is recognized for providing better geographic assignment of postal codes from rural areas that are typically linked to multiple records of standard geographic areas.[5]

4.3.2. Appendix B

Creation of income variable

We used the family after-tax income (excluding capital gains/losses) reported in the T1FF files for which the taxation years corresponded to two years prior to respondents' diagnoses. This income variable was divided by the square root of the family size to account for economies of scale in family expenditures (i.e., adjusted or equivalized family after-tax income).[6]

If the adjusted family after-tax income was missing two years prior to diagnosis, it was imputed as the average of non-missing incomes from one and three years before diagnosis. However, if income was missing one year before diagnosis, income imputation was based solely on the non-missing income from three years before diagnosis. In contrast, if income three years before diagnosis was missing or if it was missing one and three years before diagnosis, income that was missing two years before diagnosis was not imputed.

The postal codes reported in the same T1FF files were used to derive respondents' geographical location (province and urban/rural area) with the PCCF+ programs in SAS (PCCF+ versions for postal codes in years 2008-2009: 5k; and 2010-2014: 6d). In cases where postal codes were missing in the T1FF files, they were imputed by carrying forward the most recent non-missing postal codes reported for taxation years within the five years prior to the respondents' diagnoses. Each respondent's income was then matched to a population-level income tercile referenced by calendar year, province, and rurality (see below). Rurality was defined by Census Metropolitan Areas (CMAs), Census Agglomerations (CAs), rural areas, and territories, which were treated as distinct categories. We then designated income in tercile 1 as low, in tercile 2 as medium, and in tercile 3 as high.

Population-level income tercile references

Population-level income terciles were created as reference points by year, province, and rurality (CMAs: census metropolitan areas; CAs: census agglomerations; rural areas; and territories). We first linked the adjusted family after-tax income from T1FF files to all CCHS respondents (including those not part of our study) for taxation years that aligned with survey years from 2008 to 2014. For each year, we estimated survey-weighted income terciles based on geographical location (province and urban/rural area). The geographic characteristics were obtained by linking

CCHS respondents' self-reported postal codes to Statistics Canada's PCCF+ programs in SAS (PCCF+ versions for postal codes from 2008-2009: 5k; and from 2010-2014: 6d).[4] No postal codes were missing in the CCHS data. Urban/rural area was defined by CMAs, CAs, rural areas, and territories, each treated as distinct categories.

4.3.3. Appendix C

Creation of urban-rural spectrum variable

The Statistical Area Classification (SAC) type of the respondent's residence at diagnosis was used to define the urban-rural spectrum exposure. The SAC type variable categorizes census-defined geographical units, known as census subdivisions or municipalities, based on population size, density, and commuting flows. Census Metropolitan Areas (CMAs) have a population of 100,000+, of which 50,000+ reside in the core. Census Agglomerations (CAs) have a core population of 10,000+. Four types of Metropolitan Influence Zones (MIZs) consist of municipalities outside of CMAs/CAs with strong, moderate, weak, or no metropolitan influence. Territories include municipalities located beyond CAs.

The SAC type was identified using postal codes at the time of diagnosis reported in the CCR, in conjunction with PCCF+ programs in SAS (PCCF+ versions for postal codes in years 2010-2014: 6d; and 2015-2016: 7e).[4] When postal codes in the CCR were missing, we used the postal codes reported in the T1FF files for which the taxation years corresponded to respondents' diagnosis years. We consolidated SAC types into: 1) large cities (CMAs), 2) small cities (CAs), 3) rural I (MIZs with strong and moderate influence), and 4) rural II (MIZs with weak or no influence and territories).

4.3.4. Appendix D

Inclusion of survey weights

We adopted a population approach in this study by incorporating in our analyses survey weights provided by Statistics Canada that account for survey design. The weights are also adjusted to compensate for respondents who did not agree to share their survey data with Statistics Canada partners or link it with other data. The CCHS and CCR data are linked through the Derived Record Depository (DRD) in the Social Data Linkage Environment at Statistics Canada based on deterministic or probabilistic methods.[7] The survey weights are also adjusted to compensate for CCHS respondents who could not be linked to the DRD.

Following the compilation of any study sample to be used for the main analyses or secondary analyses, the survey weights from each CCHS cycle were transformed. Weights were first re-scaled using the 'pooled' approach.[8] This approach provides estimates for an average population representing residents of Canada between 2000 and 2017 who were diagnosed with lung cancer between 2010 and 2016. The first 500 bootstrap replicate weights provided by Statistics Canada for variance estimation were re-scaled using the same technique. Furthermore, all weights were normalized since many of the packages in the statistical software R that were used for our analyses do not accommodate probability weights. This normalization involved the division of survey weights by their average and of replicate weights by their sum.[9, 10] Using normalized weights ensures that the final sample size aligns with the original sample size. The sample estimated with the transformed (i.e., re-scaled and normalized) weights, denoted as the survey-weighted sample, served as the foundation for all subsequent analyses.

Table 4.3.2. Transformation of CCHS survey weights	
1. Identify the sample weights and the first 500 bootstrap weights associated with each individual observation in the final sample	Sample weight: wts_sdle Bootstrap weights: bsw1, bsw…, bsw500
2. Re-scale the survey weights with the "pooled" approach	Scaling factor: $\frac{1}{\#\ \text{CCHS cycles}} = \frac{1}{13}$ Re-scaled sample weight: wts_sdle_scale = wts_sdle * $\frac{1}{13}$ Re-scaled bootstrap weights: bsw(1…500)_scale = bsw(1…500) * $\frac{1}{13}$
3. Normalize (standardize) the re-scaled survey weights	Normalize re-scaled sample weight: wts_sdle_scale_std = $\frac{\text{wts\_sdle\_scale}}{\overline{\text{wts\_sdle\_scale\_std}}}$ Normalize re-scaled bootstrap weights: bsw(1…500)_scale_std = $\frac{\text{bsw(1…500)\_scale}}{\sum(\text{bsw(1…500)\_scale})}$

4.3.5. Appendix E

Creation of inverse-probability-of-treatment weights for survival analysis

Inverse-probability-of-treatment weights were generated and combined with transformed survey weights for survival analyses. IPTW-survey-weighted samples have covariate distributions in each exposure group that align with covariate distributions in the overall survey-weighted population.

Table 4.3.3. Estimating inverse-probability-of-treatment weights (IPTW) for survival analyses

1. Fit a weighted (multivariate) logistic regression with exposure as dependant variable and confounders (C) as independent variables	$\ln\left(\frac{P(A=1\|C)}{1-P(A=1\|C)}\right) = \beta_0 + \beta_1 C_1 + \beta_2 C_2 + \cdots + \beta_x C_x$, weights= wts_sdle_scale_std
2. Predict probability of exposure (A) conditional on confounders (C) for each observation	$P(A=1\|C) = \frac{e^{\beta_0+\beta_1 C_1+\beta_2 C_2+\cdots+\beta_x C_x}}{1+\ e^{\beta_0+\beta_1 C_1+\beta_2 C_2+\cdots+\beta_x C_x}}$ $P(A=0\|C) = 1 - \frac{e^{\beta_0+\beta_1 C_1+\beta_2 C_2+\cdots+\beta_x C_x}}{1+\ e^{\beta_0+\beta_1 C_1+\beta_2 C_2+\cdots+\beta_x C_x}}$
3. Fit a weighted (multivariate) logistic regression with exposure as dependant variable and no independent variables (i.e. intercept-only model)	$\ln\left(\frac{P(A=1)}{1-P(A=1)}\right) = \beta_0$, weights= wts_sdle_scale_std
4. Predict the marginal probability of exposure (A) for each observation	$P(A=1) = \frac{e^{\beta_0}}{1+\ e^{\beta_0}}$ $P(A=0) = 1 - \frac{e^{\beta_0}}{1+\ e^{\beta_0}}$
5. Obtain standardized inverse probability treatment weights for each observation	IPTW = $\left\lfloor (A) * \frac{P(A=1)}{P(A=1\|C)} \right\rfloor + \left\lfloor (1-\mathrm{A}) * \frac{P(A=0)}{P(A=0\|C)} \right\rfloor$
6. Combine IPTW weights with transformed weights to produce final weights.	Combined weight: wts_sdle_scale_std * IPTW Combined bootstrap weight: bsw(1…500) scale_std * IPTW
7. Incorporate combined weights in survival analyses to obtain marginal estimates	Non-parametric analysis: Kaplan-Meir (*survival* package in R), accounting for survey weights and IPTW. Semi-parametric analysis: McKeague and Sasieni's additive hazard model (*timereg* package in R), accounting for survey weights and IPTW.

4.3.6. Appendix F

Methods for mediation analyses

We applied Lange's nested counterfactual approach for mediation analyses in combination with a 'full bootstrap' procedure proposed for variance estimation in the context of survey data.[11, 12] This approach overcomes several limitations of traditional mediation analyses in the survival context. Rather than combining estimates from different outcome and mediator models that are not always collapsible, natural direct and indirect effects of exposure on outcome are estimated from a single model called the "natural effects model," which can be used with any type of outcome, mediator, and exposure, and can include exposure-mediator interactions.

Briefly, the flexible approach involves fitting an exposure model and a mediator model to the original data, simulating data, estimating weights, and using the weighted/simulated data to fit an outcome model (i.e., the natural effects model) that includes exposure-mediator interactions. For the natural effects model, we used McKeague and Sasieni's semi-parametric additive hazards model, which does not require the proportional hazards assumption. From the natural effects models, we estimated the pure natural direct effect (PNDE), the total natural indirect effect (TNIE), the total effect (TE), and the percent of the TE mediated by advanced stage at diagnosis.

PNDEs quantify the average difference in hazards due to change in exposure (e.g., switching from high to low income), while stage at diagnosis (i.e., the mediator) assumes the value it would have naturally under no exposure.

TNIEs quantify the average difference in hazards resulting from a change in the mediator from the natural value it would have taken under no exposure to the natural value under exposure, while remaining exposed.

Table 4.3.4. Estimating natural effects	
1. Using the original data, fit a weighted (multivariate) logistic regression with exposure as dependant variable and confounders (C) as independent variables	$\ln\left(\frac{P(A=1\|C)}{1-P(A=1\|C)}\right) = \beta_0 + \beta_1 C_1 + \beta_2 C_2 + \cdots + \beta_x C_x$, weights= wts_sdle_scale_std
2. Using the original data, fit a weighted logistic regression with mediator as dependant variable and exposure (A) and confounders (C) as independent variables	$\ln\left(\frac{P(M=1\|C)}{1-P(M=1\|C)}\right) = \beta_0 + \beta_1 A + \beta_2 C_1 + \beta_3 C_2 + \cdots + \beta_x C_x$, weights= wts_sdle_scale_std
3. Simulate a new dataset	Duplicate each observation* and create a new exposure variable (A'). The value of A' for the 1st observation for each ID is equal to the value of the observed A and the value of A' for the 2nd observation is equal to the counterfactual of the observed A. **The observation for each ID is duplicated as many times as the number of counterfactual exposure levels that can exist*
4. Using the simulated data, predict probability of exposure (A) conditional on confounders (C) for each observation	$P(A=1\|C) = \frac{e^{\beta_0+\beta_1 C_1+\beta_2 C_2+\cdots+\beta_x C_x}}{1+\ e^{\beta_0+\beta_1 C_1+\beta_2 C_2+\cdots+\beta_x C_x}}$ $P(A=0\|C) = 1 - \frac{e^{\beta_0+\beta_1 C_1+\beta_2 C_2+\cdots+\beta_x C_x}}{1+\ e^{\beta_0+\beta_1 C_1+\beta_2 C_2+\cdots+\beta_x C_x}}$
5. Using the simulated data, predict probability of mediator (M) conditional on exposure (A) and confounders (C) for each observation	$P(M=1\|A,C) = \frac{e^{\beta_0+\beta_1 A+\beta_2 C_1+\cdots+\beta_x C_x}}{1+\ e^{\beta_0+\beta_1 A+\beta_2 C_1+\cdots+\beta_x C_x}}$ $P(M=0\|A,C) = 1 - \frac{e^{\beta_0+\beta_1 A+\beta_2 C_1+\cdots+\beta_x C_x}}{1+\ e^{\beta_0+\beta_1 A+\beta_2 C_1+\cdots+\beta_x C_x}}$
6. Using the simulated data, predict probability of mediator (M) conditional on exposure (A') and confounders (C) for each observation	$P(M=1\|A',C) = \frac{e^{\beta_0+\beta_1 A'+\beta_2 C_1+\cdots+\beta_x C_x}}{1+\ e^{\beta_0+\beta_1 A'+\beta_2 C_1+\cdots+\beta_x C_x}}$ $P(M=0\|A',C) = 1 - \frac{e^{\beta_0+\beta_1 A'+\beta_2 C_1+\cdots+\beta_x C_x}}{1+\ e^{\beta_0+\beta_1 A'+\beta_2 C_1+\cdots+\beta_x C_x}}$
7. Obtain standardized inverse-probability-of-treatment weights (IPTW) for each observation *→ IPTW generates a pseudo treated and control samples where distribution of confounders (C) mimic the distribution in the full sample*	IPTW = $\left[(A) * \frac{P(A=1)}{P(A=1\|C)}\right] + \left[(1-\mathrm{A}) * \frac{P(A=0)}{P(A=0\|C)}\right]$
8. Obtain ratio-of-mediator-probability weights (RMPW) for each observation *→ RMPW generates a pseudo cross-world sample of treated (control) units where distribution of confounders (C) mimic the C distribution in the full sample, and the distribution of M (conditional on C) mimic the M distribution in the control (treated) units*	RMPW = $\frac{P(M=m\|A',C)}{P(M=m\|A,C)}$

Table 4.3.4 (continued). Estimating natural effects	
9. Combine IPTW and RMPW weights with transformed weights to produce final weights.	<u>Combined weight</u>: wts_sdle_scale_std * IPTW * RMPW <u>Combined bootstrap weight:</u> bsw(1…500) scale_std * IPTW * RMPW
10. Using the simulated data, incorporate final weights in mediation analyses to. Obtain marginal estimates of natural effects.	<u>Semi-parametric survival analysis:</u> Lin and Ying's additive hazard model (*timereg* package in R), accounting for survey weights, IPTW, and RMPW. $\lambda_0(t) + \beta_1 a + \beta_2 a' + \beta_3 a * a'$ Pure Natural Direct Effect (PNDE): $(\beta_1 + \beta_3 a')(a - a')$ Total Natural Indirect Effect (TNIE): $(\beta_2 + \beta_3 a)(a - a')$ Total Effect (TE): $(\beta_1 + \beta_2 + \beta_3)(a - a')$ Percent Mediated: (TNIE/TE) * 100

4.3.7. Appendix G

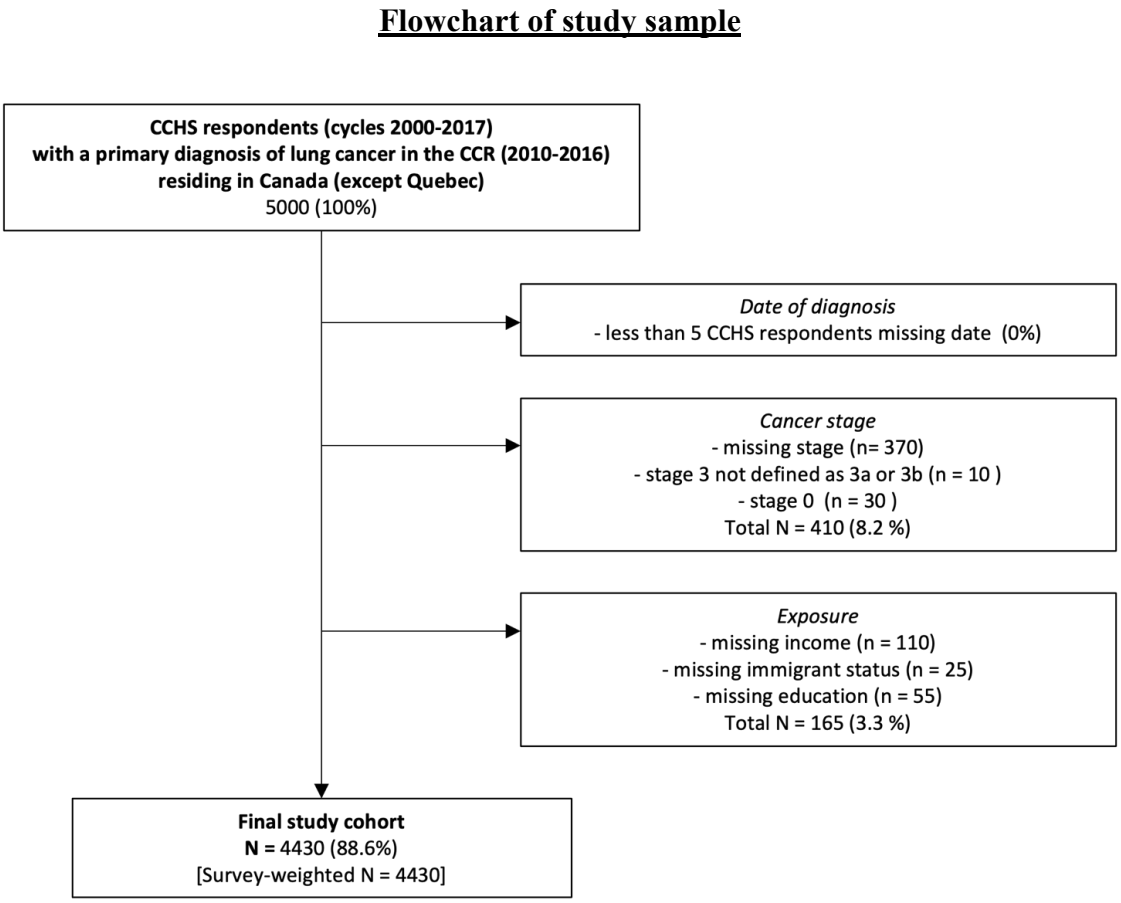


Figure 4.3.1. Flowchart of CCHS respondents included in the study

Abbreviations: CCHS: Canadian Community Health Survey; CCR: Canadian Cancer Registry.

4.3.8. Appendix H

Natural effect estimates (main, secondary, and post-hoc analyses)

Table 4.3.5. Natural effect estimates from main mediation analyses

Exposure variables	TE Hazard difference, deaths per 10,000 person-months (95% CI)	PNDE Hazard difference, deaths per 10,000 person-months (95% CI)	TNIE Hazard difference, deaths per 10,000 person-months (95% CI)	PM Percent (95% CI)
Family income terciles [A,B]				
1: Low	43.88 (-6.23, 105.48)	56.92 (14.62, 97.38)	-13.04 (-41.77, 28.69)	-29.72 (-355.87, 365.86)
2: Medium	14.12 (-33.78, 72.06)	25.58 (-18.38, 67.61)	-11.46 (-42.73, 26.54)	-81.12 (-787.73, 983.65)
3: High	REF	REF	REF	REF
Education [A]				
Less than secondary school completion	40.87 (-7.22, 80.75)	32.53 (-7.70, 66.85)	8.35 (-22.91, 35.35)	20.42 (-174.57, 221.52)
Secondary school completion and/or postsecondary education	REF	REF	REF	REF
Immigration Status [A]				
Non-immigrant (Canadian-born)	REF	REF	REF	REF
Immigrant (foreign-born)	32.56 (-31.88, 107.22)	21.88 (-35.47, 89.69)	10.59 (-25.21, 55.95)	32.79 (-453.60, 569.17)
Urban-rural spectrum [A]				
Large cities (CMAs)	REF	REF	REF	REF
Small cities (CAs)	49.11 (-4.51, 105.76)	29.13 (-9.40, 75.70)	19.98 (-11.08, 47.47)	40.69 (-74.75, 156.76)
Rural areas I (strong/moderate MIZs)	-28.50 (-72.27, 28.15)	2.50 (-39.42, 52.49)	-31.01 (-63.00, 0.78)	108.78 (-543.40, 1373.81)
Rural areas II (weak/no MIZs or territories)	-34.28 (-98.26, 27.38)	-35.13 (-84.11, 24.27)	0.85 (-29.63, 25.76)	-2.49 (-232.25, 206.59)

[A] Survey-weighted N = 4430. IPTW models used in mediation analyses accounted for age (4 categories: <65, 65-74, 75-79, and ≥80), sex, and year of diagnosis (2 categories: 2010-2013, and 2014-2016). In addition, IPTW models for income and education as exposures accounted for province (4 categories: 1- British Columbia; 2- Manitoba and Alberta; 3- New Brunswick, Newfoundland and Labrador, and Ontario; and 4- Saskatchewan, Nova Scotia, Prince Edward Island, and Territories) and rurality (3 categories: large cities, small cities, and rural areas or territories). The model for urban-rural spectrum of residence as exposure accounted for province, and the model for immigration as exposure accounted for race/ethnicity (White, and non-White, missing were imputed as White [n=20]). Provinces grouped according to previously reported 1-year survival rates.[3]

[B] Population-level terciles of equivalized after-tax family income at 2 years prior to lung cancer diagnosis, that are specific by province, urban-rural residence (CMA/CA/Rural/Territories), and calendar year.

Abbreviations: CMAs: Census Metropolitan Areas; CAs: Census Agglomerations; MIZs: Metropolitan Influence Zones; TE: total effects; PNDE: pure natural direct effects; TNIE: total natural indirect effects; PM: percent mediated.

Table 4.3.6. Natural effect estimates from secondary mediation analyses

Exposure variables	TE Hazard difference, deaths per 10,000 person-months (95% CI)	PNDE Hazard difference, deaths per 10,000 person-months (95% CI)	TNIE Hazard difference, deaths per 10,000 person-months (95% CI)	PM Percent (95% CI)
A. CCHS respondents from waves 2000 to 2014, survey-weighted N = 4140				
Family income terciles [A, B]				
1: Low	51.28 (-6.57, 119.45)	58.09 (8.74, 102.67)	-6.82 (-38.98, 38.29)	-13.29 (-217.71, 279.08)
2: Medium	15.42 (-40.43, 79.17)	25.12 (-22.41, 75.49)	-9.70 (-44.42, 31.13)	-62.92 (-670.85, 812.71)
3: High	REF	REF	REF	REF
Education [A]				
Less than secondary school completion	54.29 (-0.85, 101.45)	42.39 (-0.94, 78.98)	11.90 (-23.47, 41.46)	21.91 (-161.98, 105.31)
Secondary school completion and/or postsecondary education	REF	REF	REF	REF
Immigration status [A]				
Non-immigrant (Canadian-born)	REF	REF	REF	REF
Immigrant (foreign-born)	27.73 (-39.05, 103.40)	17.34 (-47.16, 87.64)	10.39 (-28.05, 61.17)	37.47 (-279.27, 797.58)
Urban-rural spectrum [A]				
Large cities (CMAs)	REF	REF	REF	REF
Small cities (CAs)	60.10 (1.49, 122.98)	32.51 (-14.50, 83.37)	27.59 (-9.69, 59.31)	45.91 (-53.67, 149.13)
Rural areas I (strong/moderate MIZs)	-30.20 (-79.04, 29.35)	-0.14 (-46.40, 52.77)	-30.06 (-64.03, 1.87)	99.55 (-628.80, 860.45)
Rural areas II (weak/no MIZs or territories)	-44.98 (-110.58, 22.73)	-50.72 (-104.52, 14.10)	5.74 (-27.15, 31.87)	-12.76 (-189.20, 341.53)
B. CCHS respondents from waves 2000 to 2016 diagnosed with lung cancer after CCHS participation, survey-weighted N = 4010				
Family income terciles [A, B]				
1: Low	62.14 (-6.41, 134.47)	60.46 (3.40, 113.74)	1.68 (-37.60, 47.40)	2.71 (-241.81, 140.54)
2: Medium	36.62 (-31.16, 101.39)	44.62 (-12.57, 101.55)	-8.00 (-47.23, 38.40)	-21.85 (-525.46, 313.09)
3: High	REF	REF	REF	REF
Education [A]				
Less than secondary school completion	42.75 (-15.05, 96.64)	35.35 (-14.95, 78.44)	7.40 (-30.95, 41.43)	17.31 (-312.80, 413.96)
Secondary school completion and/or postsecondary education	REF	REF	REF	REF
Immigration status [A]				
Non-immigrant (Canadian-born)	REF	REF	REF	REF
Immigrant (foreign-born)	8.60 (-72.60, 99.75)	2.64 (-64.65, 78.53)	5.96 (-37.05, 56.37)	69.31 (-639.63, 703.83)
Urban-rural spectrum [A]				
Large cities (CMAs)	REF	REF	REF	REF
Small cities (CAs)	64.00 (2.32, 133.33)	40.39 (-10.13, 99.31)	23.61 (-14.65, 60.91)	36.89 (-62.17, 147.21)
Rural areas I (strong/moderate MIZs)	-22.37 (-80.16, 41.87)	8.86 (-44.07, 67.83)	-31.23 (-67.81, 6.67)	139.61 (-1978.54, 1101.13)
Rural areas II (weak/no MIZs or territories)	-60.38 (-129.90, 14.68)	-54.72 (-116.48, 7.98)	-5.67 (-42.44, 25.37)	9.38 (-107.02, 177.93)

Table 4.3.6 (continued). Natural effect estimates from secondary mediation analyses

Exposure variables	TE Hazard difference, deaths per 10,000 person-months (95% CI)	PNDE Hazard difference, deaths per 10,000 person-months (95% CI)	TNIE Hazard difference, deaths per 10,000 person-months (95% CI)	PM Percent (95% CI)
C. CCHS respondents from waves 2000 to 2017; advanced stage (mediator) redefined as stages IIIa, IIIb, and IV; survey-weighted N = 4430				
Family income terciles [A, B]				
1: Low	44.10 (-6.15, 105.40)	50.45 (9.18, 90.62)	-6.35 (-39.44, 37.37)	-14.40 (-429.63, 268.43)
2: Medium	14.34 (-33.72, 72.46)	26.89 (-15.05, 67.03)	-12.55 (-43.40, 24.37)	-87.55 (-1342.38, 1049.53)
3: High	REF	REF	REF	REF
Education [A]				
Less than secondary school completion	40.85 (-7.31, 80.58)	21.44 (-16.40, 55.80)	19.41 (-8.24, 45.14)	47.52 (-107.67, 283.31)
Secondary school completion and/or postsecondary education	REF	REF	REF	REF
Immigration status [A]				
Non-immigrant (Canadian-born)	REF	REF	REF	REF
Immigrant (foreign-born)	32.97 (-31.92, 107.66)	39.88 (-14.83, 89.87)	-6.91 (-41.78, 47.16)	-20.94 (-831.73, 630.52)
Urban-rural spectrum [A]				
Large cities (CMAs)	REF	REF	REF	REF
Small cities (CAs)	47.90 (-5.18, 104.61)	1.38 (-33.90, 46.71)	46.52 (12.72, 74.73)	97.12 (-116.74, 401.39)
Rural areas I (strong/moderate MIZs)	-29.18 (-72.04, 27.67)	-25.66 (-62.43, 22.77)	-3.53 (-35.62, 25.60)	12.08 (-362.43, 528.94)
Rural areas II (weak/no MIZs or territories)	-34.80 (-97.61, 26.13)	-47.93 (-97.54, 7.91)	13.13 (-18.92, 40.61)	-37.72 (-440.68, 246.80)
D. CCHS respondents from waves 2000 to 2017; males; survey-weighted N = 1955				
Family income terciles [A, B]				
1: Low	2.04 (-84.10, 95.22)	46.86 (-25.21, 112.83)	-44.82 (-104.25, 20.98)	-2191.65 (-1113.76, 2256.15)
2: Medium	21.29 (-78.39, 98.19)	53.85 (-43.28, 131.62)	-32.56 (-97.00, 32.99)	-152.97 (-887.95, 1110.71)
3: High	REF	REF	REF	REF
Education [A]				
Less than secondary school completion	47.61 (-35.50, 114.04)	31.80 (-32.45, 93.29)	15.81 (-37.33, 58.92)	33.21 (-371.26, 270.71)
Secondary school completion and/or postsecondary education	REF	REF	REF	REF
Immigration status [A]				
Non-immigrant (Canadian-born)	REF	REF	REF	REF
Immigrant (foreign-born)	-20.78 (-110.83, 83.89)	-20.81 (-111.48, 77.60)	0.03 (-57.38, 66.54)	-0.14 (-551.41, 972.74)
Urban-rural spectrum [A]				
Large cities (CMAs)	REF	REF	REF	REF
Small cities (CAs)	77.27 (-13.53, 176.34)	61.23 (-15.23, 149.10)	16.04 (-50.17, 74.29)	20.76 (-200.45, 267.02)
Rural areas I (strong/moderate MIZs)	-46.86 (-117.69, 42.31)	-8.63 (-77.77, 77.78)	-38.23 (-102.40, 11.04)	81.57 (-457.74, 854.25)
Rural areas II (weak/no MIZs or territories)	-14.96 (-147.02, 118.47)	-3.75 (-99.10, 111.59)	-11.21 (-79.94, 35.10)	74.93 (-464.39, 888.15)

Table 4.3.6 (continued). Natural effect estimates from secondary mediation analyses

Exposure variables	TE Hazard difference, deaths per 10,000 person-months (95% CI)	PNDE Hazard difference, deaths per 10,000 person-months (95% CI)	TNIE Hazard difference, deaths per 10,000 person-months (95% CI)	PM Percent (95% CI)
E. CCHS respondents from waves 2000 to 2017; females; survey-weighted N = 2475				
Family income terciles [A, B]				
1: Low	77.53 (17.47, 142.17)	67.83 (17.32, 127.22)	9.69 (-25.41, 51.96)	12.51 (-111.57, 58.64)
2: Medium	19.14 (-38.22, 84.40)	19.45 (-27.47, 72.43)	-0.30 (-34.21, 39.87)	-1.59 (-313.93, 643.69)
3: High	REF	REF	REF	REF
Education [A]				
Less than secondary school completion	32.02 (-22.59, 86.26)	29.63 (-16.71, 71.04)	2.39 (-30.90, 39.02)	7.47 (-338.82, 389.28)
Secondary school completion and/or postsecondary education	REF	REF	REF	REF
Immigration status [A]				
Non-immigrant (Canadian-born)	REF	REF	REF	REF
Immigrant (foreign-born)	78.66 (-2.50, 176.97)	51.56 (-18.27, 128.18)	27.10 (-23.41, 78.55)	34.45 (-71.83, 157.51)
Urban-rural spectrum [A]				
Large cities (CMAs)	REF	REF	REF	REF
Small cities (CAs)	32.13 (-26.81, 85.50)	12.61 (-29.13, 60.49)	19.53 (-11.75, 48.82)	60.77 (-169.19, 390.59)
Rural areas I (strong/moderate MIZs)	-6.62 (-56.92, 55.73)	21.39 (-27.49, 79.68)	-28.01 (-57.69, 6.53)	423.21 (-1280.57, 1708.96)
Rural areas II (weak/no MIZs or territories)	-37.69 (-94.81, 29.08)	-43.69 (-93.82, 16.83)	6.00 (-21.29, 36.98)	-15.91 (-335.09, 390.67)

[A] IPTW models used in mediation analyses accounted for age (4 categories: <65, 65-74, 75-79, and ≥80), sex (when applicable), and year of diagnosis (2 categories: 2010-2013, and 2014-2016). In addition, IPTW models for income and education as exposures accounted for province (4 categories: 1- British Columbia; 2- Manitoba and Alberta; 3- New Brunswick, Newfoundland and Labrador, and Ontario; and 4- Saskatchewan, Nova Scotia, Prince Edward Island, and Territories) and rurality (3 categories: large cities, small cities, and rural areas or territories). The model for urban-rural spectrum of residence as exposure accounted for province, and the model for immigration as exposure accounted for race/ethnicity (White and non-White, missing were imputed as White [n=20]). Provinces grouped according to previously reported 1-year survival rates.[3]

[B] Population-level terciles of equivalized after-tax family income at 2 years prior to lung cancer diagnosis, that are specific by province, urban-rural residence (CMA/CA/Rural/Territories), and calendar year.

Abbreviations: CCHS: Canadian Community Health Survey; CMAs: Census Metropolitan Areas; CAs: Census Agglomerations; MIZs: Metropolitan Influence Zones; TE: total effects; PNDE: pure natural direct effects; TNIE: total natural indirect effects; PM: percent mediated.

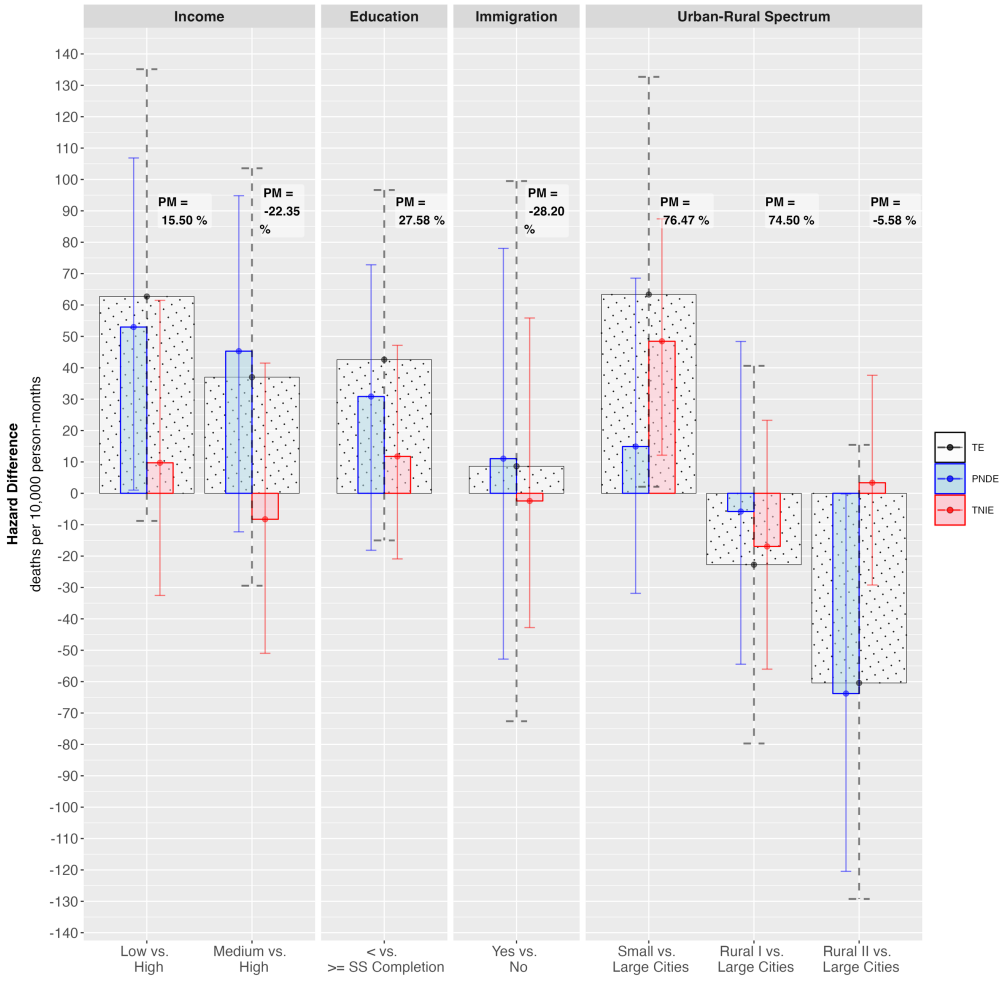


Figure 4.3.2. Post-hoc mediation analyses for income, education, immigration status, and urban-rural spectrum, including survey, IPTW, and mediation weights.

CCHS respondents from waves 2000 to 2016 diagnosed with lung cancer after CCHS participation; advanced stage (mediator) redefined as stages IIIa, IIIb, and IV; survey-weighted N = 4010 .

The grey dotted bars represent total effects, the blue bars represent the pure natural direct effects, the red bars represent the total natural indirect effects, and the corresponding whiskers represent the 95% confidence intervals.

IPTW models used in mediation analyses accounted for age (4 categories: <65, 65-74, 75-79, and ≥80), sex, and year of diagnosis (2 categories: 2010-2013, and 2014-2016). In addition, IPTW models for income and education as exposures accounted for province (4 categories: 1- British Columbia; 2- Manitoba and Alberta; 3- New Brunswick, Newfoundland and Labrador, and Ontario; and 4- Saskatchewan, Nova Scotia, Prince Edward Island, and Territories) and rurality (3 categories: large cities, small cities, and rural areas or territories). The model for urban-rural spectrum of residence as exposure accounted for province, and the model for immigration as exposure accounted for race/ethnicity (White and non-White, missing were imputed as White [n=20]). Provinces grouped according to previously reported 1-year survival rates.[3]

Abbreviations: SS: secondary school; TE: total effects; PNDE: pure natural direct effects; TNIE: total natural indirect effects; PM: percent mediated.

Table 4.3.7. Natural effect estimates from post-hoc mediation analyses

Exposure variables	TE Hazard difference, deaths per 10,000 person-months (95% CI)	PNDE Hazard difference, deaths per 10,000 person-months (95% CI)	TNIE Hazard difference, deaths per 10,000 person-months (95% CI)	PM Percent (95% CI)
CCHS respondents from waves 2000 to 2016 diagnosed with lung cancer after CCHS participation; advanced stage (mediator) redefined as stages IIIa, IIIb, and IV; survey-weighted N = 4010				
Family income terciles [A, B]				
1: Low	62.70 (-8.80, 135.13)	52.98 (1.01, 106.85)	9.72 (-32.54, 61.49)	15.50 (-181.99, 175.68)
2: Medium	37.01 (-29.44, 103.57)	45.29 (-12.27, 94.83)	-8.27 (-50.98, 41.49)	-22.35 (-620.61, 456.26)
3: High	REF	REF	REF	REF
Education [A]				
Less than secondary school completion	42.59 (-15.01, 96.64)	30.84 (-18.15, 72.82)	11.75 (-20.90, 47.15)	27.58 (-196.79, 242.00)
Secondary school completion and/or postsecondary education	REF	REF	REF	REF
Immigration status [A]				
Non-immigrant (Canadian-born)	REF	REF	REF	REF
Immigrant (foreign-born)	8.62 (-72.62, 99.49)	11.05 (-52.83, 78.04)	-2.43 (-42.78, 55.86)	-28.20 (-650.00, 589.86)
Urban-rural spectrum [A]				
Large cities (CMAs)	REF	REF	REF	REF
Small cities (CAs)	63.34 (2.09, 132.69)	14.91 (-31.87, 68.55)	48.44 (12.14, 87.45)	76.47 (-17.36, 260.42)
Rural areas I (strong/moderate MIZs)	-22.73 (-79.72, 40.62)	-5.80 (-54.46, 48.39)	-16.93 (-56.03, 23.26)	74.50 (-1631.92, 964.98)
Rural areas II (weak/no MIZs or territories)	-60.43 (-129.25, 15.44)	-63.80 (-120.45, -0.19)	3.37 (-29.22, 37.61)	-5.58 (-142.06, 298.96)

[A] IPTW models used in mediation analyses accounted for age (4 categories: <65, 65-74, 75-79, and ≥80), sex, and year of diagnosis (2 categories: 2010-2013, and 2014-2016). In addition, IPTW models for income and education as exposures accounted for province (4 categories: 1- British Columbia; 2- Manitoba and Alberta; 3- New Brunswick, Newfoundland and Labrador, and Ontario; and 4- Saskatchewan, Nova Scotia, Prince Edward Island, and Territories) and rurality (3 categories: large cities, small cities, and rural areas or territories). The model for urban-rural spectrum of residence as exposure accounted for province, and the model for immigration as exposure accounted for race/ethnicity (White and non-White, missing were imputed as White [n=20]). Provinces grouped according to previously reported 1-year survival rates.[3]

[B] Population-level terciles of equivalized after-tax family income at 2 years prior to lung cancer diagnosis, that are specific by province, urban-rural residence (CMA/CA/Rural/Territories), and calendar year.

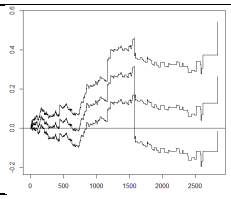
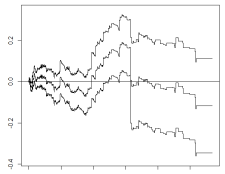
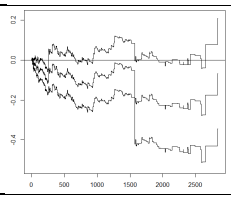
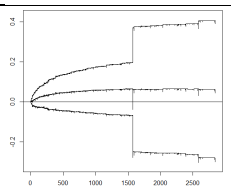
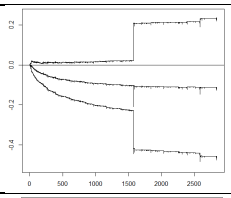
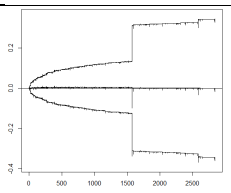
Abbreviations: CCHS: Canadian Community Health Survey; CMAs: Census Metropolitan Areas; CAs: Census Agglomerations; MIZs: Metropolitan Influence Zones; TE: total effects; PNDE: pure natural direct effects; TNIE: total natural indirect effects; PM: percent mediated.

4.3.9. Appendix I

Assessing time-varying natural effect estimates for urban-rural spectrum

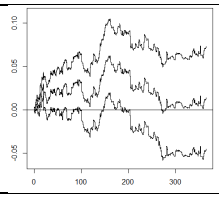
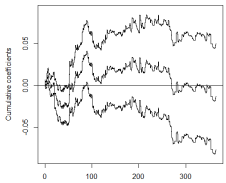
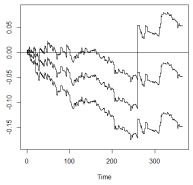
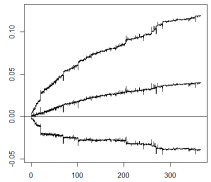
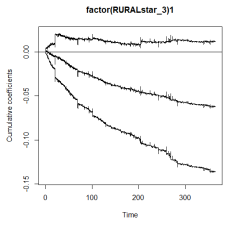
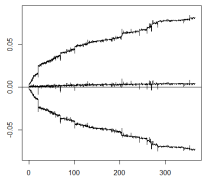
Step 1. We fit Aalen's nonparametric additive hazard model using the same weighted and simulated dataset in the main analysis for the urban-rural spectrum of residence, but without any terms capturing exposure-mediator interaction. Plots and goodness-of-fit tests did not take into account CCHS bootstrap weights. Survey-weighted N = 4430.

Table 4.3.8. Assessment of time-varying natural effects for urban-rural spectrum

Natural effect	Contrast [reference: large cities]	Plot of cumulative regression coefficients against time (days) with 95% confidence intervals	Cramer von Mises test p-value (H_0=constant effect)	Kolmogorov-Smirnov test p-value (H_0=constant effect)
Direct	Small cities		0.526	0.529
	Rural areas I		0.236	0.107
	Rural areas II		0.370	0.426
Indirect	Small cities		0.000	0.422
	Rural 1		0.000	0.123
	Rural 2		0.316	0.276

Step 2. Step 1 was repeated, but with follow-up times restricted to 12 months. Survey-weighted N = 4430.

Table 4.3.9. Assessment of time-varying natural effects for urban-rural spectrum, within 12 months of follow-up

Natural effect	Contrast [reference: large cities]	Plot of cumulative regression coefficients against time (days) with 95% confidence intervals	Cramer von Mises test p-value (H_0=constant effect)	Kolmogorov-Smirnov test p-value (H_0=constant effect)
Direct	Small cities		0.413	0.272
	Rural 1		0.500	0.646
	Rural 2		0.387	0.315
Indirect	Small cities		0.000	0.311
	Rural I		0.000	0.076
	Rural II		0.160	0.778

Step 3. We fit McKeague and Sasieni's semi-parametric additive hazard model (modelling constant hazard differences) using the same weighted and simulated dataset as in the main analysis for the urban-rural spectrum of residence, but without any terms capturing exposure-mediator interaction and with follow-up times restricted to 12 months. Survey-weighted N = 4430.

Table 4.3.10. Natural effect estimates from mediation analyses with follow-up time limited to 12 months

Exposure variable	TE Hazard difference, deaths per 10,000 person-months (95% CI)	PNDE Hazard difference, deaths per 10,000 person-months (95% CI)	TNIE Hazard difference, deaths per 10,000 person-months (95% CI)	PM Percent (95% CI)
Urban-rural spectrum [A]				
Large cities (CMAs)	REF	REF	REF	REF
Small cities (CAs)	59.68 (-41.25, 179.86)	21.49 (-62.60, 122.38)	38.19 (-20.23, 91.18)	63.99 (-418.72, 477.42)
Rural areas I (strong/moderate MIZs)	-66.89 (-166.19, 44.01)	-1.75 (-84.38, 95.80)	-65.13 (-128.89, 1.41)	97.38 (-723.58, 1132.55)
Rural areas II (weak/no MIZs or territories)	-49.44 (-182.31, 62.68)	-52.59 (-164.21, 45.59)	3.14 (-60.74, 54.92)	-6.35 (-443.28, 282.96)

[A] IPTW models used in mediation analyses accounted for age (4 categories: <65, 65-74, 75-79, and ≥80), sex, and year of diagnosis (2 categories: 2010-2013, and 2014-2016). In addition, IPTW models for income and education as exposures accounted for province (4 categories: 1- British Columbia; 2- Manitoba and Alberta; 3- New Brunswick, Newfoundland and Labrador, and Ontario; and 4- Saskatchewan, Nova Scotia, Prince Edward Island, and Territories) and rurality (3 categories: large cities, small cities, and rural areas or territories). The model for urban-rural spectrum of residence as exposure accounted for province, and the model for immigration as exposure accounted for race/ethnicity (White and non-White, missing were imputed as White [n=20]). Provinces grouped according to previously reported 1-year survival rates.[3]

Abbreviations: CCHS: Canadian Community Health Survey; CMAs: Census Metropolitan Areas; CAs: Census Agglomerations; MIZs: Metropolitan Influence Zones; TE: total effects; PNDE: pure natural direct effects; TNIE: total natural indirect effects; PM: percent mediated.

References for appendices

(1) Yusuf W, Vyuha R, Bennett C, Sequeira Y, Maskerine C, Manuel DG. cchsflow: an open science approach to transform and combine population health surveys. Canadian Journal of Public Health. 2021;112:714-21.

(2) Johnson CH, Peace S, Adamo P, Fritz A, Percy-Laurry A, Edwards BK. The 2007 multiple primary and histology coding rules. Bethesda, MD: National Cancer Institute, Surveillance, Epidemiology and End Results Program; 2007.

(3) Canadian Cancer Society's Advisory Committee. Canadian Cancer Statistics: A 2020 special report on lung cancer. Toronto, ON: Canadian Cancer Society; 2020.

(4) Postal Code OM Conversion File Plus (PCCF+) [Internet]. Government of Canada. 2023 [cited 2023-12-06]. Available from: https://www150.statcan.gc.ca/n1/en/catalogue/82F0086X#wb-auto-2.

(5) Wilkins R. PCCF+ Version 5E User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through March 2009. In: Health Analysis Division, editor. Ottawa, ON.2009.

(6) Pinard D. Methodology Changes: Census Family Low Income Measure Based on the T1 Family File: Statistics Canada; 2018.

(7) Social Data Linkage Environment (SDLE) [Internet]. Government of Canada. 2022. Available from: https://www.statcan.gc.ca/en/sdle/index.

(8) Thomas S, Wannell B. Combining cycles of the Canadian community health survey. Health reports. 2009;20(1):53.

(9) Rao J, Wu C, Yue K. Some recent work on resampling methods for complex surveys. Survey methodology. 1992;18(2):209-17.

(10) Gagné C, Roberts G, Keown L-A. Weighted estimation and bootstrap variance estimation for analyzing survey data: How to implement in selected software. The Research Data Centres Information and Technical Bulletin. 2014;6(1):1-72.

(11) Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. Am J Epidemiol. 2012;176(3):190-5. DOI: 10.1093/aje/kwr525

(12) Reaume M, McIsaac MA. Natural direct and indirect effects for survival outcomes in a complex survey setting. Biostatistics Practicum Report, Queen's University, Kingston, Canada. 2017.

Chapter 5. Manuscript 2

5.1. Preface

Lung cancer patients with advanced-stage disease face a universally poor prognosis, regardless of social position. This is due to limited treatment options and the disease's inherently aggressive and fatal nature. However, recent advancements in precision medicine, such as the introduction of targeted therapies, are gradually shifting this paradigm by improving survival outcomes for advanced lung cancer patients. As patients begin to survive longer, social inequalities in survival may emerge. Access to and timely initiation of these breakthrough therapies is often a function of social and geographic factors and, therefore may not be equitable. In the era of targeted therapies for advanced lung cancer, overall improved survival outcomes may come with increased social inequalities in survival.

For the study in this manuscript, I leverage Quebec's health administrative data, which captures prescriptions filled in community pharmacies, to identify users of targeted therapies. By linking this cohort to other health administrative data, I investigate survival differences based on SES, sociodemographic factors, and place of residence. Additionally, I explore disparities in the use and timely initiation of these therapies across the same populations. Building on the findings from the previous manuscript, this study emphasizes the importance of addressing lung cancer survival inequalities that extend beyond differences due to stage at diagnosis. With more than three-fourths of lung cancer cases diagnosed at an advanced stage, this study examines how social factors contribute to survival disparities at these later stages of the disease.

This manuscript was published in the peer-reviewed journal *Cancer Medicine*:

Qureshi S, Boily G, Boulanger J, Pagé É, Strumpf E. Inequalities in survival and care across social determinants of health in a cohort of advanced lung cancer patients in Quebec (Canada): A high-resolution population-level analysis. *Cancer Med*. 2023;12(11):12683-704. DOI: 10.1002/cam4.5897

5.2. Inequalities in survival and care across social determinants of health in a cohort of advanced lung cancer patients in Quebec (Canada): a high-resolution population-level analysis

Samia Qureshi[1*], Gino Boily[2], Jim Boulanger[3], Élisabeth Pagé[2], Erin Sttumpf[1]

[1]Department of Epidemiology and Biostatistics, McGill University, H3A 1G1, Montreal, Quebec, Canada

[2] Institut national d'excellence en santé et services sociaux (INESSS), H3A 2S9, Montreal, Quebec, Canada

[3] Institut national d'excellence en santé et services sociaux (INESSS), H3A 2S9, Quebec, Quebec, Canada

[*]**Corresponding Author:** Samia Qureshi

Address for Correspondence: samia.qureshi@mail.mcgill.ca

Key Words: health inequalities, lung neoplasms, molecular targeted therapy, sociodemographic factors, socioeconomic factors, survival analysis, time-to-treatment

Word Count: Abstract: 249**;** Main text: 3,451

Abstract

Background: Advanced lung cancer patients exposed to breakthrough therapies like EGFR tyrosine kinase inhibitors (EGFR-TKI) may experience social inequalities in survival, partly from differences in care. This study examined survival by neighbourhood-level socioeconomic and sociodemographic status, and geographical location of advanced lung cancer patients who received gefitinib, an EGFR-TKI, as 1st-line palliative treatment. Differences in the use and delay of EGFR-TKI treatment were also examined.

Methods: Lung cancer patients receiving gefitinib from 2001 to 2019 were identified from Quebec's health administrative databases. Accounting for age and sex, estimates were obtained for the median time of survival from the time of treatment to death, the probability of receiving osimertinib as a 2nd EGFR-TKI, and the median time from biopsy to receiving 1st-line gefitinib.

Results: Among 457 patients who received 1st-line treatment with gefitinib, those living in the most materially deprived areas had the shortest median survival time (high/low deprivation, 0.69; 95% CI, 0.47-1.04). The probability of receiving osimertinib as a 2nd EGFR-TKI was highest for patients from immigrant-dense areas (high/low density, 1.95; 95% CI, 1.26-3.36) or from Montreal (other urban areas/Montreal: 0.39; 95% CI, 0.16-0.71). The median wait time for gefitinib was 1.27 times longer in regions with health centers peripheral to large centers in Quebec or Montreal in comparison to regions with university-affiliated centers (95% CI, 1.09-1.54; n=353).

Conclusion: This study shows that real-world variations in survival and treatment exist among advanced lung cancer patients in the era of breakthrough therapies and that future research on inequalities should also focus on population.

Introduction

Lung cancer remains the deadliest cancer in Canada.[1] Roughly, half of all lung cancer patients are diagnosed at stage IV, when the 1-year survival probability is as little as 17%.[1, 2] With the gradual introduction of breakthrough palliative treatments like precision medicine (i.e., targeted therapies in 2010 and immunotherapy in 2015),[3, 4] survival trends are expected to rise in the overall and advanced (i.e., locally advanced or metastatic) lung cancer populations.[5, 6] However, survival improvements in socioeconomically, sociodemographically, or geographically disadvantaged subpopulations may not necessarily occur at the same pace as in advantaged subpopulations. Despite a universal healthcare system providing free health services at point of care for medically necessary treatments across all provinces, there is growing evidence for social inequalities in outcomes along the cancer control continuum in Canada.[7]

Due to suboptimal care, survival from different cancers, including lung cancer, may be shorter in socioeconomically, sociodemographically, or geographically disadvantaged groups.[8-11] For example, cancer patients with low socioeconomic status (SES), immigrant status, or living in specific regions in Canada are less likely to receive oncology consultations, surgery, radiotherapy, chemotherapy, and supportive care near death.[12-15] Cancer treatment delays can also impact survival,[16] including delays for targeted treatment in advanced lung cancer.[17] However, the evidence for SES- and geography-based differences in timely treatment for lung cancer in Canada and elsewhere is mixed.[14, 18-21]

There are very few lung cancer studies in Canada that focus on inequalities in survival by SES measures other than income or by sociodemographic status (SDS) (e.g. immigration), and on inequalities in palliative care with systemic treatments.[1, 13] Also, income-based inequalities in survival are not evident for Canadian patients diagnosed with stage IV lung cancer in 2010-2011,[1] a period when precision medicine barely existed. Decision-makers in Quebec are increasingly interested in the real-world effectiveness of breakthrough palliative therapies for advanced lung cancer, like EGFR tyrosine kinase inhibitors (EGFR-TKI). While the survival of EGFR-TKI users in Quebec has been investigated,[22, 23] inequalities in survival and treatments by patient subgroups remain unknown. Investigating these inequalities in Quebec is relevant, since Quebec accounts for 23% of the Canadian population [24] and about 16% of all Canadian immigrants in Canada.[25] Among the most populous provinces (Ontario, Quebec, British

Columbia, and Alberta), it has the lowest median household income (Quebec vs. others in 2020 constant Canadian dollars: $53,800 vs. $63,000-$75,300) [26] and a higher percentage of individuals with educational attainment below upper secondary levels (Quebec vs. others: 12% vs. 7-8%).[27] Studies of advanced lung cancer patients in the United States suggest that there may be SES-based inequalities in survival,[28] in the use of breakthrough palliative treatments,[29] and geography-based inequalities in the delay to initiate EGFR-TKI treatment.[30]

This population-based retrospective study sought to fill a knowledge gap on survival inequalities and potential drivers of these inequalities across understudied characteristics of advanced lung cancer patients living in the current context of breakthrough therapies in Canada. Among patients who received the first-generation of EGFR-TKIs, gefitinib, as 1st-line palliative treatment in Quebec, we examined fluctuations in survival by income, education, material deprivation index, immigration, and geographic region, as well as variations in the use and delay of EGFR-TKI treatment (i.e., use of a 2nd EGFR-TKI, osimertinib, and delay between the tumour biopsy and 1st EGFR-TKI). Results from this study may also be relevant to other countries with similar socio-economic contexts, such as developed or high-income countries (i.e., US, UK, European Union, and Australia) that have social inequalities in lung cancer treatment and survival,[9, 10, 12, 19, 20] and countries with universal healthcare systems which include coverage of oral cancer medicines.[31]

Material and methods

Study design and data

We conducted a population-level retrospective study by selecting patients in Quebec whose first indication of lung cancer was between April 1, 2001, and March 31, 2019, and who received gefitinib as 1st-line palliative treatment in that period. Data came from health administrative databases managed by the Régie de l'assurance maladie du Québec (RAMQ), that capture information related to public insurance plans for health and prescription drugs, and by the Ministère de la Santé et des Services sociaux (MSSS), that capture inpatient and outpatient health care utilization by the Quebec population. Gefitinib is only reimbursed in Quebec to treat advanced lung cancer patients who are naïve to systemic palliative treatment. Focusing on this treatment thus allowed us to identify advanced lung cancer patients that are close in time to their diagnosis date.

The dates for the first indication of lung cancer and the initial lung cancer diagnosis were determined based on previously established methods.[32] Patients who filled a prescription for gefitinib at least once between April 1, 2001 and March 31, 2019 were identified in the community pharmacy services database through previously selected international nonproprietary names (INN) and drug identification numbers (DIN).[22, 23] To form cohort 1, which was used to study survival and the use of osimertinib as a 2nd EGFR-TKI, we identified the line of treatment associated with patients' 1st gefitinib treatment using a verified algorithm [22, 23] and excluded all patients who did not receive gefitinib as a 1st-line palliative treatment. From this cohort (cohort 1), a separate cohort (cohort 2) was formed to study the time delay from receiving a tumour biopsy (or biopsy-related intervention) to receiving treatment. To form cohort 2, we excluded patients who did not receive a biopsy or received it more than 6 weeks before their lung cancer diagnosis and/or more than 6 months before their 1st gefitinib treatment. The date of death was captured with the health insurance registry (FIPA).

This study was part of a Health System Impact Fellowship of the Canadian Institutes of Health Research and was conducted at the Institut national d'excellence en santé et services sociaux (INESSS). De-identified data was accessible through a tripartite agreement between MSSS, RAMQ, and INESSS.[33] INESSS is not responsible for the content of this publication, however, the cohort used in this study is also included in a larger study by INESSS on EGFR-TKIs in Quebec.[22, 23]

Outcomes

Survival time was calculated in months from the date of 1st gefitinib prescription to either the date of death or March 31, 2020: the administrative censoring time. Previously selected INNs and DINs were used to identify patients who received osimertinib as a 2nd EGFR-TKI during their follow-up after receiving gefitinib.[22, 23] Medical interventions involving biopsies that occurred between patients' diagnosis dates,[32] and their 1st gefitinib treatments were extracted from physician billing and hospitalization databases with expert-validated billing and intervention codes (Appendix A).[34, 35] Treatment delays were calculated as the number of days between each patient's most recent biopsy and their 1st gefitinib treatment.

Exposures and covariates

FIPA was consulted for patients' age, sex, and areas of residence (1: the census-based dissemination area [DA], which are the smallest standard geographic unit [400-700 persons] at the national level; and 2: one of 18 health regions) at the time of their 1st gefitinib treatment. Patients' DAs were linked to neighbourhood-level SES and SDS, and geography variables developed by INESSS with the 2016 census data and the overall RAMQ population in that year. The DA-level SES and SDS variables included categorical information on median income after tax, percent of low education, and percent of individuals who immigrated to Canada. The census-based geography variable classified patients' DA based on urbanicity: 1) DAs in Census Metropolitan Areas (CMA), which include Montreal and other large cities, were urban; 2) DAs in Census Agglomerations (CA) were suburban; and 3) all other DAs (Non-CMAs/CAs) were rural. Montreal was analyzed separately from other CMAs since it includes 86% of all immigrants in Quebec.[36] Patients DAs were also linked to Pampalon's material deprivation index, which was developed with Quebec's 2016 Census population and is a composite score of income, education, and employment measures at the DA-level that have been standardized to Quebec's age and sex distributions.[37] Another health center-based geography variable we used delineated patients' health regions according to the presence of university-affiliated medical centers and the distance of other health centers from large health services centres in Montreal and Quebec.[38, 39] More details on each exposure are listed in Appendix B (Table 5.3.2). Comorbidities were calculated with the Population Grouping Methodology,[40, 41] which uses diagnostic codes to search for 226 health conditions (excluding lung cancer) in the three years leading up to the date of the 1st gefitinib treatment.

Statistical analyses

Medians and proportions were calculated for patients' baseline characteristics. The median follow-up time in cohort 1 was estimated with the distribution of censoring times (reverse Kaplan-Meier method) with the *prodlim* package in R.[42] We used the *survival* package in R to conduct inverse-probability-of-treatment weighted (IPTW) Kaplan-Meier analyses.[43] Stabilized weights for IPTW were calculated with predicted treatment probabilities obtained from multinomial logistic regressions. The weighing produced a pseudo-population with age and sex distributions at each category of SES, SDS, and geography that are like those in the overall unweighted population. For each exposure level, survival curves were plotted. The ratios of median overall survival (OS) times

were also estimated from the Kaplan-Meier analyses, and 95% confidence intervals were obtained with a stratified case resampling method for right-censored data.[44]

Osimertinib, a third-generation EGFR-TKI, can increase overall survival when prescribed in place of chemotherapy as a subsequent line after another EGFR-TKI.[45] We set out to identify variations in the use of this treatment, which might explain potential survival inequalities observed in the same cohort. Therefore, we ran age- and sex-adjusted quasipoisson regressions to estimate the ratios of marginal probabilities of receiving osimertinib as a 2nd EGFR-TKI treatment across SES, SDS, and geography in cohort 1. Bootstrapping was used to obtain 95% confidence intervals. Osimertinib was first approved for coverage by Quebec's public drug insurance program in November 2018. We assumed that the fraction of patients in our study that died before this period, who could not receive osimertinib, was non-differential across SES, SDS, and geography. The validity of this assumption was tested in a secondary analysis that included only patients who were alive after October 31, 2018.

In cohort 2, we estimated the 25th, 50th, and 75th percentiles of treatment delays as the shorter, median, and longer delays, respectively, in each category of SES, SDS, and geography. The weighting method applied to cohort 1 was also used for this analysis. At each percentile (25th, 50th, or 75th), we estimated ratios of delays across categories of SES, SDS, and geography, and obtained bootstrapped 95% confidence intervals.

With an expected small sample, our exploratory analyses did not include statistical significance testing. We interpreted our results based on the direction and magnitude of the point estimates and the width of the associated confidence intervals.

Results

Cohort characteristics

Of the 552 patients who received gefitinib between April 1, 2001, and March 31, 2019, 457 patients (82.8%) received gefitinib as a 1st-line palliative treatment, so they were included in cohort 1. Of the 457 patients, 353 (77.2%) patients were included in cohort 2 (Figure 5.2.1). Only 1 patient received their 1st gefitinib before the fiscal year 2011, and more than half of all patients received it in 2015 or afterwards, both in cohort 1 [23] and cohort 2. For patients in cohort 1, the median age at 1st gefitinib treatment was 70 years, and the median time to treatment from diagnosis was 3

months (IQR: 1.9-12.9 months) (Table 5.2.1). A large percentage of patients were female (68.5%) or did not receive osimertinib as a 2nd EGFR-TKI treatment (79.9%). Most patients had between 0 and 9 comorbidities (73%), of which half had less than 5 comorbidities (36.5%). A higher proportion of patients resided in areas with the lowest income quintile (26.0%), or the highest quintiles of low education (25.4%) or immigration density (31.9%). At least half of all patients resided in Montreal (55.6%) or in areas with university-affiliated health centers (51.0%). There were no major differences in cohort 2 in comparison to cohort 1, except for a slightly shorter time from diagnosis to treatment (median: 2.6 months, IQR: 1.8-4.9 months). Because of the small sizes of the middle categories of income, education, and immigration variables, and the suburban category of the census-based region variable, we refrained from interpreting any comparisons made with these categories.

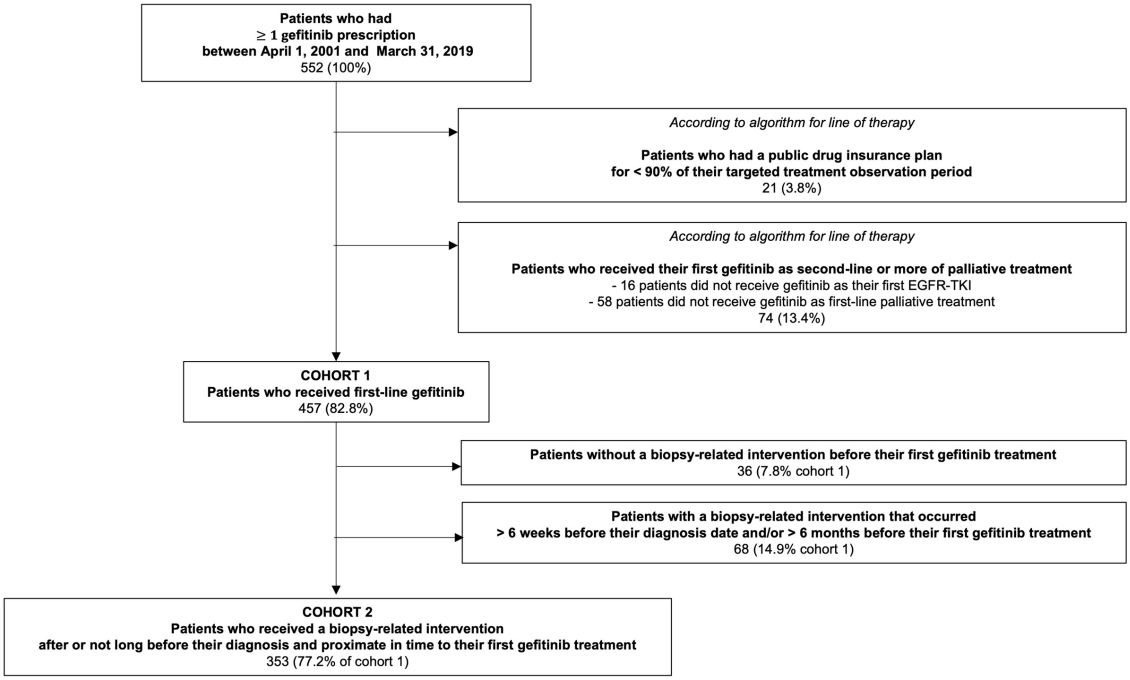


Figure 5.2.1. Flowchart of study population in each cohort

Table 5.2.1. Patient characteristics (unweighted populations)

	Cohort #1 n (%)	**Cohort #2 n (%)**
Age in Years		
< 60	58 (12.7)	51 (14.4)
60-64	48 (10.5)	43 (12.2)
65-70	102 (22.3)	82 (23.2)
≥ 70	249 (54.5)	177 (50.1)
Median (IQR)	70.6 (65.4 - 77.6)	70.0 (64.2 - 77.1)
Female	313 (68.5 %)	252 (71.3 %)
Comorbidities		
0-4	169 (37.0)	129 (36.5)
5-9	160 (35.0)	129 (36.5)
≥ 10	127 (27.7)	94 (26.6)
Missing	1 (0.2)	1 (0.28)
Median (IQR)	6 (3 - 10)	6 (3 - 10)
Osimertinib as 2nd EGFR-TKI Treatment	92 (20.1)	73 (20.7)
1st osimertinib in 2016	10 (2.2)	10 (2.8)
1st osimertinib in 2017	20 (4.4)	14 (4.0)
1st osimertinib in 2018	42 (9.2)	34 (9.6)
1st osimertinib in 2019	20 (4.4)	15 (4.2)
Months Between Diagnosis and 1st Gefitinib Treatment Median (IQR)	3 (1.9 - 12.9)	2.6 (1.8 - 4.9)
Median Income After Tax (Quintiles)		
Low (Quintiles 1 and 2): 0 - 28,222$	200 (43.7)	150 (42.4)
Medium (Quintile 3): 28,223$ - 31,653$	99 (21.7)	76 (21.5)
High (Quintiles 4 and 5): ≥ 31,654$	151 (33.1)	123 (34.9)
Missing	7 (1.5)	7 (1.1)
Percent with Low Education (Quintiles)		
Low (Quintiles 1 and 2): 0 - 28.69 %	206 (45.1)	166 (46.9)
Medium (Quintile 3): 28.7 - 34.99 %	77 (16.8)	54 (15.3)
High (Quintiles 4 and 5): ≥ 35.00 %	171 (37.5)	132 (37.4)
Missing	3 (0.66)	3 (0.28)
Material Deprivation Index (Terciles)		
Low (Tercile 1: Least Deprived)	147 (32.2)	117 (33.1)
Medium (Tercile 2)	126 (27.6)	99 (28.0)
High (Tercile 3: Most Deprived)	152 (33.3)	114 (32.3)
Missing	32 (7.0)	23 (6.5)

Table 5.2.1. Patient characteristics (unweighted populations)

	Cohort #1 n (%)	**Cohort #2 n (%)**
Percent Immigrant (Quintiles)		
Low (Quintiles 1 and 2): 0 - 4.09 %	152 (33.3)	114 (32.4)
Medium (Quintile 3): 4.10 - 10.29 %	70 (15.3)	57 (16.1)
High (Quintiles 4 and 5): ≥ 10.30 %	232 (50.7)	71 (51.3)
Missing	3 (0.66)	1 (0.28)
Census-Based Region		
Urban: Montreal	254 (55.6)	194 (55.0)
Urban: Others	83 (18.2)	65 (18.4)
Suburban	41 (9.0)	33 (9.3)
Rural	77 (16.8)	60 (16.9)
Missing	2 (0.44)	1 (0.28)
Health Center-Based Region		
University	233 (51.0)	178 (50.4)
Peripheral	142 (31.1)	116 (32.9)
Intermediary	64 (14.0)	48 (13.6)
Remote or Northern	16 (3.5)	10 (2.8)
Missing	2 (4.4)	1 (0.28)
Total (N)	**457**	**353**

Abbreviations: IQR: interquartile range

Overall survival

With a total of 337 deaths, the median follow-up and overall survival times were respectively 45.3 months (IQR: 27.3-57.1) and 19.8 months (95% CI: 17.1-23.1 months). Our application of IPTW was successful in balancing age and sex distributions across categories of SES, SDS, and geographic location (Appendix C). Age- and sex-standardized survival patterns across SES, SDS, and geographic location in the weighted populations are presented in Figure 5.2.2. The ratios in median survival times in Table 5.2.2 that were obtained from the inverse probability-weighted survival curves show that patients from neighbourhoods with the highest percentage of low education had lower median survival time (high/low ratio: 0.82, 95% CI: 0.64-1.17), as did those living in areas with the highest level of material deprivation index (high/low ratio: 0.69, 95% CI: 0.47-1.04). In contrast, patients from neighbourhoods with the highest immigrant density had longer survival (high/low ratio: 1.23, 95% CI: 0.84-1.73). Patients living in urban areas other than Montreal had shorter survival than those living in Montreal (time ratio: 0.71, 95% CI: 0.52-1.04). The ratios of survival by income groups and health center-based regions were the smallest.

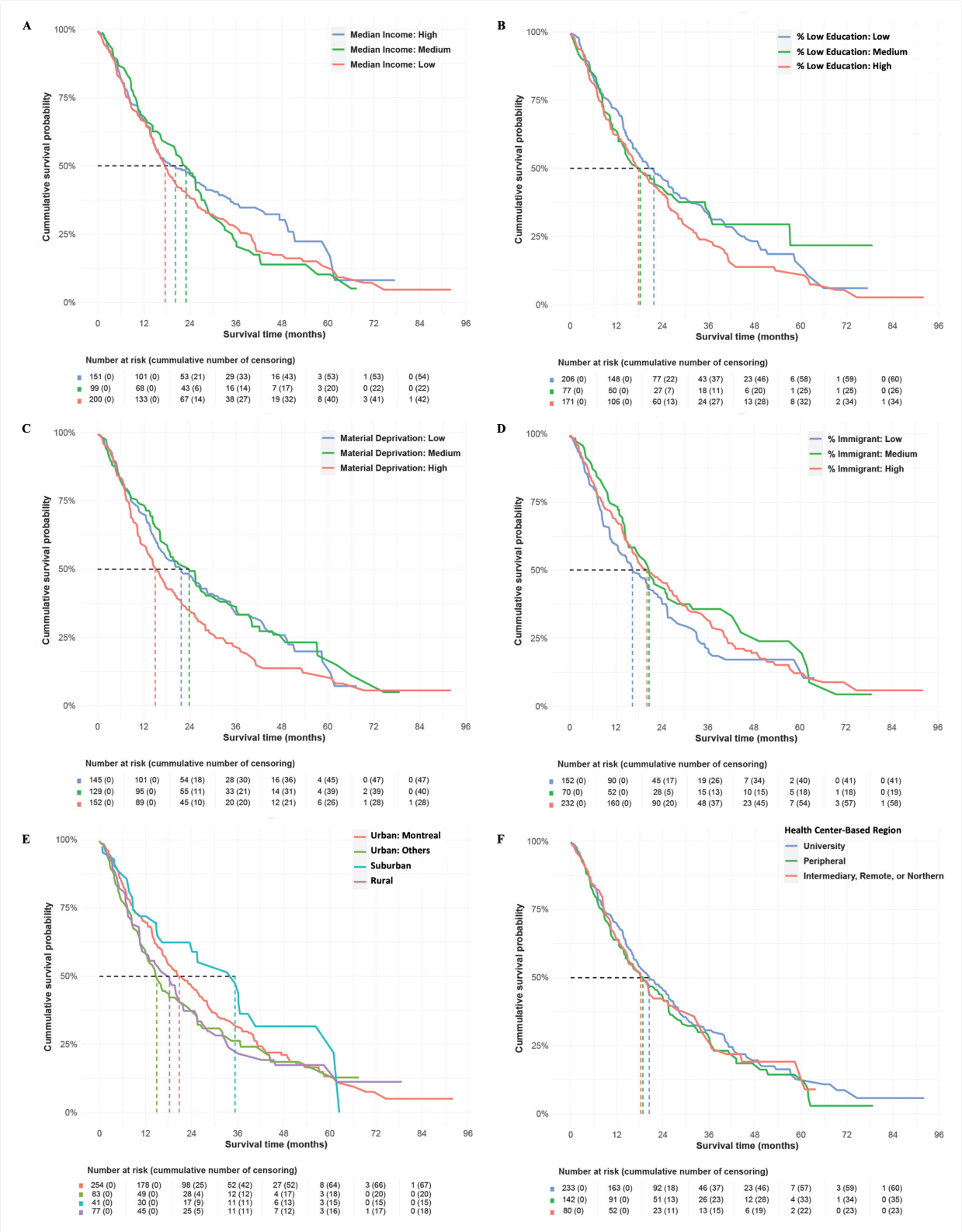


Figure 5.2.2. Overall survival with gefitinib as first-line treatment, by socioeconomic, sociodemographic, and geographic factors (age- and sex-standardized, weighted populations)

Table 5.2.2. Median overall survival (age- and sex-standardized, weighted populations)

	Median Survival Time in Months[1] (95% CI)	Median Survival Time Ratio[1] (95% CI)
SES and SDS Variables		
Income		
Low: 0 - 28,222$	17.6 (15.1 - 21.1)	0.87 (0.60 - 1.26)
Medium: 28,223$ - 31,653$	23.0 (16.7 - 27.0)	1.14 (0.71 - 1.66)
High: ≥ 31,654$	20.2 (14.6 - 28.0)	REF
Low Education		
Low: 0 - 28.69 %	21.7 (17.3 - 25.6)	REF
Medium: 28.7 - 34.99 %	18.3 (12.6 - 28.3)	0.84 (0.56 - 1.49)
High: ≥ 35.00 %	17.8 (15.0 - 23.0)	0.82 (0.64 - 1.17)
Material Deprivation Index		
Low (Least Deprived)	21.7 (16.4 - 29.6)	REF
Medium	23.8 (18.1 - 29.4)	1.10 (0.70 - 1.59)
High (Most Deprived)	15.1 (11.8 - 18.2)	0.69 (0.47 - 1.04)
Immigration		
Low: 0 - 4.09 %	16.4 (12.6 - 22.9)	REF
Medium: 4.10 - 10.29 %	20.7 (14.6 - 28.3)	1.26 (0.77 - 1.93)
High: ≥ 10.30 %	20.1 (16.3 - 25.9)	1.23 (0.84 - 1.73)
Geography Variables		
Census-Based Region		
Urban: Montreal	20.8 (17.6 - 25.9)	REF
Urban: Others	14.9 (11.9 - 21.7)	0.71 (0.52 - 1.04)
Suburban	35.3 (15.1 - 40.6)	1.70 (0.69 - 2.16)
Rural	18.2 (11.1 - 21.9)	0.87 (0.51 - 1.16)
Health Center-Based Region		
University	20.5 (16.7 - 25.5)	REF
Peripheral	18.8 (14.3 - 24.1)	0.92 (0.63 - 1.31)
Intermediary, Remote or Northern	18.3 (13.1 - 32.1)	0.89 (0.57 - 1.58)

Abbreviations: SES: socioeconomic status; SDS: sociodemographic status. [1]Analyses accounted for age (<60, 60-64, 65-70, ≥70 years) and sex.

Probability of receiving osimertinib as 2nd EGFR-TKI treatment

Estimates in Table 5.2.3 show that patients living in neighbourhoods with the highest levels of low education, or material deprivation index, in comparison to the lowest levels, had lower probabilities of receiving osimertinib as a 2nd EGFR-TKI treatment (high/low ratio and 95% CI, respectively: 0.69, 0.43-1.11; and 0.75, 0.43-1.22). Patients from the highest immigrant-dense areas were twice as likely to receive osimertinib (high/low ratio: 1.95, 95% CI: 1.26-3.36). Conversely, patients living in other urban areas or rural areas had much lower probabilities of receiving osimertinib compared to patients from Montreal (percentage ratio and 95% CI,

respectively: 0.39, 0.16-0.71; and 0.55, 0.25-0.98). The ratios across income and health center-based regions were the smallest.

Table 5.2.3. Receiving osimertinib as a 2nd EGFR-TKI after receiving gefitinib (age and sex-standardized, unweighted populations)

	Percentage[1] (95% CI)	Percentage Ratio[1] (95% CI)
SES and SDS Variables		
Income		
Low: 0 - 28,222$	17.8 (12.5 - 22.6)	0.85 (0.54 - 1.33)
Medium: 28,223$ - 31,653$	22.2 (14.4 - 30.4)	1.05 (0.62 - 1.72)
High: ≥ 31,654$	21.5 (14.7 - 28.8)	REF
Low Education		
Low: 0 - 28.69 %	24.3 (18.6 - 29.5)	REF
Medium: 28.7 - 34.99 %	15.8 (7.6 - 24.2)	0.65 (0.31 - 1.14)
High: ≥ 35.00 %	16.9 (11.1 - 22.8)	0.69 (0.43 - 1.11)
Material Deprivation Index		
Low (Least Deprived)	21.7 (15.6 - 29.1)	REF
Medium	24.2 (16.8 - 32.9)	1.12 (0.69 - 1.82)
High (Most Deprived)	16.3 (10.3 - 22.2)	0.75 (0.43 - 1.22)
Immigration		
Low: 0 - 4.09 %	12.7 (7.4 - 18.1)	REF
Medium: 4.10 - 10.29 %	19.8 (11.0 - 28.8)	1.56 (0.75 - 3.0)
High: ≥ 10.30 %	24.8 (19.4 - 29.9)	1.95 (1.26 - 3.36)
Geography Variables		
Census-Based Region		
Urban: Montreal	24.1 (18.5 - 29.6)	REF
Urban: Others	9.4 (3.8 - 16.1)	0.39 (0.16 - 0.71)
Suburban	28.9 (14.7 - 44.5)	1.20 (0.60 - 2.05)
Rural	13.3 (5.9 - 22.1)	0.55 (0.25 - 0.98)
Health Center-Based Region		
University	19.8 (14.4 - 25.4)	REF
Peripheral	22.7 (15.3 - 29.0)	1.15 (0.74 - 1.71)
Intermediary, Remote or Northern	16.1 (8.1 - 24.9)	0.81 (0.42 - 1.38)

[1]Analyses accounted for age (<60, 60-64, 65-70, ≥70 years) and sex.

Abbreviations: SES: socioeconomic status; SDS: sociodemographic status.

Treatment delays

The overall median time between a biopsy and 1st gefitinib treatment was 39 days (IQR: 26-55). The largest ratios of treatment delays were observed by geographical location and across all quantiles of treatment delays in the weighted populations (Table 5.2.4). Patients living in other urban areas waited shorter than those in Montreal (delay ratios and 95% CIs of 25th, 50th, and 75th percentiles, respectively: 0.70, 0.56-1.07; 0.79, 0.66-0.95; and 0.83, 0.63-1.00). In contrast, patients living in peripheral regions waited longer than those in university regions (delay ratios and 95% CIs of 25th, 50th, and 75th percentiles, respectively: 1.48, 1.18-2.07; 1.27, 1.09-1.54; and 1.29, 1.11-1.48).

Table 5.2.4. Delay in receiving 1st-line treatment after receiving a biopsy-related intervention (age and sex-standardized, weighted populations)

	25th Percentile		50th Percentile		75th Percentile	
	Delay in Days[1] (95% CI)	**Delay Ratio[1] (95% CI)**	**Delay in Days[1] (95% CI)**	**Delay Ratio[1] (95% CI)**	**Delay in Days[1] (95% CI)**	**Delay Ratio[1] (95% CI)**
SES and SDS Variables						
Income						
Low: 0 - 28,222$	25.3 (20.4 - 31.9)	1.05 (0.77 - 1.49)	39 (36 - 42)	1.08 (0.86 - 1.21)	52 (47 - 59.9)	0.90 (0.76 - 1.09)
Medium: 28,223$ - 31,653$	27.8 (22.7 - 32.7)	1.15 (0.91 - 1.45)	37.5 (32.6 - 42)	1.04 (0.82 - 1.23)	51.0 (43 - 57)	0.89 (0.73 - 1.08)
High: ≥ 31,654$	24 (19.7 - 29.4)	REF	36 (33 - 42.2)	REF	57.5 (49.7 - 65)	REF
Low Education						
Low: 0 - 28.69 %	23.2 (19.6 - 27.1)	REF	36 (33.5 - 40.9)	REF	55 (48.8 - 60.4)	REF
Medium: 28.7 - 34.99 %	30 (25.7 - 36)	1.29 (1.06 - 1.74)	40.5 (35.8 - 45)	1.13 (0.96 - 1.27)	54.6 (45 - 64.8)	0.99 (0.80 - 1.17)
High: ≥ 35.00 %	28 (22 - 33)	1.21 (0.96 - 1.58)	40 (35 - 42)	1.11 (0.93 - 1.22)	54.2 (48 - 61.9)	0.99 (0.86 - 1.15)
Material Deprivation Index						
Low (Least Deprived)	21 (19.0 - 28.8)	REF	36 (33 - 45.8)	REF	57 (50 - 66)	REF
Medium	28.8 (27 - 32.2)	1.37 (1.00 - 1.58)	39 (35 - 44)	1.08 (0.88 - 1.25)	52 (46 - 59)	0.91 (0.75 - 1.08)
High (Most Deprived)	26.2 (21.8 - 32.8)	1.25 (0.87 - 1.57)	39 (35 - 42)	1.08 (0.85 - 1.22)	51.5 (45 - 58)	0.90 (0.76 - 1.06)
Immigration						
Low: 0 - 4.09 %	26 (21 - 28)	REF	35 (30 - 40)	REF	48.7 (42 - 57.5)	REF
Medium: 4.10 - 10.29 %	29.0 (21 - 36.1)	1.11 (0.80 - 1.49)	40 (35 - 47.3)	1.14 (0.92 - 1.41)	55.9 (49 - 73.8)	1.15 (0.91 - 1.55)
High: ≥ 10.30 %	25.0 (20.0 - 31)	0.96 (0.77 - 1.27)	39 (36 - 43)	1.11 (0.92 - 1.32)	56.0 (50.0 - 62.4)	1.15 (0.93 - 1.36)
Geography Variables						
Census-Based Region						
Urban: Montreal	30 (23.0 - 33)	REF	42 (38 - 46.8)	REF	58 (53.9 - 64.2)	REF
Urban: Others	21 (17 - 29.9)	0.70 (0.56 - 1.07)	33 (29.2 - 37.9)	0.79 (0.66 - 0.95)	48 (37 - 58)	0.83 (0.63 - 1.00)
Suburban	27.0 (21.9 - 34.0)	0.90 (0.71 - 1.26)	38.8 (28 - 45)	0.92 (0.63 - 1.08)	48.9 (40 - 88)	0.84 (0.66 - 1.53)
Rural	23.4 (21 - 29)	0.78 (0.66 - 1.13)	35 (27.8 - 41)	0.83 (0.65 - 1.02)	52.2 (40.6 - 67	0.90 (0.67 - 1.13)
Health Center-Based Region						
University	21 (15.5 - 25)	REF	35 (31 - 37.6)	REF	48 (44.6 - 53.3)	REF
Peripheral	31 (27.0 - 35)	1.48 (1.18 - 2.07)	44.5 (40 - 54)	1.27 (1.09 - 1.54)	62 (56.9 - 70)	1.29 (1.11 - 1.48)
Intermediary, Remote or Northern	27.3 (22.2 - 32.8)	1.30 (1.03 - 1.81)	39 (30.5 - 42)	1.11 (0.86 - 1.29)	51.9 (42 - 65)	1.08 (0.81 - 1.34)

Abbreviations: SES: Socioeconomic status; SDS: Sociodemographic status. [1]Analyses accounted for age (<60, 60-64, 65-70, ≥70 years) and sex.

Discussion

Using population-based health-administrative data, we investigated whether inequalities in survival and treatment services exist among advanced lung cancer patients receiving gefitinib as 1st-line palliative treatment in Quebec. We observed shorter survival among patients residing in areas most representative of low education, high material deprivation, and minimal immigrant density in relation to the least representative areas. Patients from rural areas, and from urban areas other than Montreal, had similarly shorter median survival times than Montreal patients.

Our conclusions are qualitatively like several studies in Canada and elsewhere that investigated inequalities in survival from lung cancer by income, immigration status, and rurality in stage non-specific populations.[14, 28, 46-53] Two Canadian studies report no associations of survival with income in advanced lung cancer patients or with rurality in lung cancer patients of whom at least 50% were diagnosed at stage IV.[1, 50] However, these studies were conducted in a period where breakthrough treatments were not available or in a province with inequities in public coverage of oral versus intravenous cancer drugs (e.g., Ontario).[54]

The inequalities we observed in the probability of receiving osimertinib as a 2nd EGFR-TKI were in the same direction as our survival inequalities. Given that, across material deprivation, we observed the largest variation in survival, but the smallest in receiving osimertinib, it is unlikely that longer survival solely explains the larger variations in osimertinib use observed across census-based regions, immigration, and education. In our secondary analysis, patients residing in intermediary, remote or northern health regions had a considerably lower likelihood of receiving osimertinib in comparison to university regions (Appendix D), confirming the possibility of confounding by calendar time in our main health region-based analyses. Our results on the use of an EGFR-TKI (i.e. osimertinib) across income are similar to a previous study in Ontario reporting a small positive relationship.[14]

The inequalities in delay from the time of biopsy to 1st-line targeted treatment we observed were not in line with the survival inequalities. Montreal patients experienced longer delays but also longer survival than those from other census-based regions, and delays by health center-based regions were the largest despite lacking important survival differences. It is possible that there is a threshold effect and that the observed contrasts in treatment delays (e.g., 45 vs 35 days) do not

translate into meaningful survival differences. The overall median delay of 39 days we observed is in line with delays reported in other studies in Quebec: 1) a median of 36 days from biopsy to 1st targeted therapy;[55] 2) a median of 29 days from diagnosis to 1st palliative treatment,[56] and; 3) ≤ 28 days from diagnosis to 1st targeted therapy for 66% of patients receiving it.[57] In our study, all medians of treatment delay were above 30 days, with the highest being 45 days for peripheral regions. While our estimates are close to targets adopted by Quebec institutions of 28 and 42 days from diagnosis to 1st lung cancer treatment,[55-57] the inter-regional variation highlights room for improvement. Institutions in Canada have also demonstrated the achievability of delays as low as 7-15 days from the 1st visit involving a pathological diagnosis of advanced lung cancer to targeted treatment initiation.[58, 59] Since the 25th, 50th, and 75th percentiles of treatment delay in peripheral health regions were consistently 10 days longer than those in university regions, we believe that the health center-based gap in treatment delays is more likely due to health system factors than patients' disease severity.

One limitation of this study is the small sample size which resulted in reduced statistical power and widening the confidence intervals. However, this is inevitable since approximately 12% of advanced lung cancer patients receive EGFR-TKIs in Quebec.[60] Our study was restricted to patients with public drug insurance, who may be older than privately insured patients. This may induce a selection bias given that access to private insurance is driven by income and employment, which are impacted by health and age, factors that also affect survival. However, we assume a low proportion of privately insured patients in Quebec among all gefitinib users during the study period. Patients aged 65 and over represent 70% of all new lung cancer cases [32, 61] and belong to a provincial-level age group in which 90% of residents are registered for public drug coverage.[54, 62] Our analysis also accounted for differences in age across the different income groups, which allowed us to partially control for potential selection bias pathways diluting the relationship between SES and survival. Even though Canada's universal public health care system does not include outpatient prescription drugs, Quebec has a universal drug insurance policy that requires all residents to have prescription drug coverage (residents without private coverage are covered by a public program). Furthermore, oral cancer drugs that are available on the public formulary are fully covered by the public program.[54] It would be interesting to repeat our analyses with health administrative data from other provinces in Canada that have a low potential for selection bias due to drug insurance. For example, Alberta Manitoba, and Saskatchewan, offer

universal public coverage of cancer drugs that are available on the public formulary with a full reimbursement, and have only 12-13% of all take-home cancer drugs sales paid by private insurance.[63] Lastly, we estimated associations that cannot be explained by differences in age and sex, however, future analyses should jointly account for potential confounders like immigration (or geographical location), comorbidities, smoking, ethnicity, and calendar time.

To our knowledge, this is the first study investigating SES-, SDS-, and geography-based inequalities in survival and treatment in a cohort of advanced lung cancer patients in Canada, in the era of breakthrough therapies. Our analyses suggest that socioeconomic, sociodemographic, and geographic inequalities in survival and in care with EGFR-TKIs exist among patients receiving gefitinib as 1st-line palliative treatment. Given that the use of breakthrough therapies is expanding and that about half of all patients are diagnosed at an advanced stage, we expect to see survival inequalities over time, in a similar direction as in our study, in the advanced lung cancer population. Future studies on survival inequalities should therefore include this population in a causal framework to identify mediators that can be acted upon with policy and clinical interventions.

Funding

This research was part of a doctoral Health System Impact Fellowship awarded to S.Q. The fellowship was co-funded by the Canadian Institutes of Health Research (CIHR-Institutes of Health Services and Policy Research and Cancer Research) and the Institut national d'excellence en santé et services sociaux (funding number HI9-166408). E.S.'s is also a holder of McGill University's William Dawson Scholar award, which supports her research.

Conflict of interest

The authors declare no other conflicts of interest.

Ethics approval

Ethics approval was also granted by McGill University's Faculty of Medicine and Health Sciences Institutional Review Board on March 22, 2022 (study number: A04-M31-22A).

Author contributions

Samia Qureshi: Conceptualization, methodology, validation, formal analysis, data curation, writing—original draft preparation, writing—review and editing, visualization, funding acquisition **Gino Boily:** Methodology, validation, writing—review and editing **Jim Boulanger:** Methodology, validation, writing—review and editing **Élisabeth Pagé:** Resources, writing—review and editing, supervision **Erin Strumpf:** Conceptualization, methodology, validation, writing—review and editing, supervision

Data availability

The data that support the findings of this study are available on request from the Institut national d'excellence en santé et services sociaux (INESSS). The data are not publicly available due to privacy or ethical restrictions.

References

(1) Canadian Cancer Society's Advisory Committee. Canadian Cancer Statistics: A 2020 special report on lung cancer. Toronto, ON: Canadian Cancer Society; 2020.

(2) Walters S, Maringe C, Coleman MP, et al. Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: a population-based study, 2004-2007. Thorax. 2013;68(6):551-64. DOI: 10.1136/thoraxjnl-2012-202297

(3) Ellis PM, Verma S, Sehdev S, Younus J, Leighl NB. Challenges to implementation of an epidermal growth factor receptor testing strategy for non-small-cell lung cancer in a publicly funded health care system. J Thorac Oncol. 2013;8(9):1136-41. DOI: 10.1097/JTO.0b013e31829f6a43

(4) Juergens RA, Mariano C, Jolivet J, et al. Real-world benefit of nivolumab in a Canadian non-small-cell lung cancer cohort. Curr Oncol. 2018;25(6):384-92. DOI: 10.3747/co.25.4287

(5) Ten Berge D, Aarts MJ, Groen HJM, Aerts J, Kloover JS. A population-based study describing characteristics, survival and the effect of TKI treatment on patients with EGFR mutated stage IV NSCLC in the Netherlands. Eur J Cancer. 2022;165:195-204. DOI: 10.1016/j.ejca.2022.01.038

(6) Howlader N, Forjaz G, Mooradian MJ, et al. The Effect of Advances in Lung-Cancer Treatment on Population Mortality. N Engl J Med. 2020;383(7):640-9. DOI: 10.1056/NEJMoa1916623

(7) Lambert LK, Horrill TC, Beck SM, et al. Health and healthcare equity within the Canadian cancer care sector: a rapid scoping review. Int J Equity Health. 2023;22(1):20. DOI: 10.1186/s12939-023-01829-2

(8) Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer survival: a review. Ann Oncol. 2006;17(1):5-19. DOI: https://doi.org/10.1093/annonc/mdj007

(9) Carriere R, Adam R, Fielding S, Barlas R, Ong Y, Murchie P. Rural dwellers are less likely to survive cancer – An international review and meta-analysis. Health & Place. 2018;53:219-27. DOI: https://doi.org/10.1016/j.healthplace.2018.08.010

(10) Afshar N, English DR, Milne RL. Rural–urban residence and cancer survival in high-income countries: A systematic review. Cancer. 2019;125(13):2172-84. DOI: https://doi.org/10.1002/cncr.32073

(11) Ahmed S, Shahid RK. Disparity in cancer care: a Canadian perspective. Curr Oncol. 2012;19(6):e376-e82. DOI: 10.3747/co.19.1177

(12) Redondo-Sánchez D, Petrova D, Rodríguez-Barranco M, Fernández-Navarro P, Jiménez-Moleón JJ, Sánchez MJ. Socio-Economic Inequalities in Lung Cancer Outcomes: An

Overview of Systematic Reviews. Cancers (Basel). 2022;14(2). DOI: 10.3390/cancers14020398

(13) Maddison AR, Asada Y, Urquhart R. Inequity in access to cancer care: a review of the Canadian literature. Cancer Causes Control. 2011;22(3):359-66. DOI: 10.1007/s10552-010-9722-3

(14) Shah M, Parmar A, Chan KKW. Socioeconomic disparity trends in diagnostic imaging, treatments, and survival for non-small cell lung cancer 2007-2016. Cancer medicine. 2020;9(10):3407-16. DOI: https://doi.org/10.1002/cam4.2978

(15) Chu A, Barbera L, Sutradhar R, Oz UE, O'Leary E, Seow H. Association between end-of-life cancer care and immigrant status: a retrospective cohort study in Ontario, Canada. BMJ Open. 2021;11(6):e042978. DOI: 10.1136/bmjopen-2020-042978

(16) Hanna TP, King WD, Thibodeau S, et al. Mortality due to cancer treatment delay: systematic review and meta-analysis. BMJ. 2020;371:m4087. DOI: 10.1136/bmj.m4087

(17) Sheinson D, Wong WB, Wu N, Mansfield AS. Impact of delaying initiation of anaplastic lymphoma kinase inhibitor treatment on survival in patients with advanced non-small-cell lung cancer. Lung Cancer. 2020;143:86-92. DOI: 10.1016/j.lungcan.2020.03.005

(18) Muslim Z, Stroever S, Razi SS, et al. Increasing Time-to-Treatment for Lung Cancer: Are We Going Backward? The Annals of Thoracic Surgery. 2022. DOI: https://doi.org/10.1016/j.athoracsur.2022.06.016

(19) Forrest LF, Sowden S, Rubin G, White M, Adams J. Socio-economic inequalities in stage at diagnosis, and in time intervals on the lung cancer pathway from first symptom to treatment: systematic review and meta-analysis. Thorax. 2017;72(5):430. DOI: https://doi.org/10.1136/thoraxjnl-2016-209013

(20) Jacobsen MM, Silverstein SC, Quinn M, et al. Timeliness of access to lung cancer diagnosis and treatment: A scoping literature review. Lung Cancer. 2017;112:156-64. DOI: 10.1016/j.lungcan.2017.08.011

(21) Campbell NC, Elliott AM, Sharp L, Ritchie LD, Cassidy J, Little J. Impact of deprivation and rural residence on treatment of colorectal and lung cancer. Br J Cancer. 2002;87(6):585-90. DOI: 10.1038/sj.bjc.6600515

(22) Institut national d'excellence en santé et en services sociaux (INESSS). Utilisation en contexte québécois des inhibiteurs de la tyrosine kinase du récepteur du facteur de croissance épidermique (EGFR) pour le traitement du cancer du poumon. État des pratiques rédigé par Gino Boily, Kossi Thomas Golo, Aude- Christine Guédon, Camille Lehuédé, Samia Qureshi, Ferdaous Roussafi et Erin Strumpf. Québec, Qc2022. p. 109.

(23) Qureshi S, Boily G, Boulanger J, et al. Advanced Lung Cancer Patients' Use of EGFR Tyrosine Kinase Inhibitors and Overall Survival: Real-World Evidence from Quebec, Canada. Curr Oncol. 2022;29(11):8043-73. DOI: 10.3390/curroncol29110636

(24) Statistics Canada. Table 17-10-0005-01 Population estimates on July 1st, by age and sex. 2022.

(25) Statistics Canada. Table 17-10-0008-01 Estimates of the components of demographic growth, annual. 2022.

(26) Statistics Canada. Table 11-10-0190-01 Market income, government transfers, total income, income tax and after-tax income by economic family type. 2022.

(27) Statistics Canada. Table 37-10-0130-01 Educational attainment of the population aged 25 to 64, by age group and sex, Organisation for Economic Co-operation and Development (OECD), Canada, provinces and territories. 2022.

28) Maguire FB, Morris CR, Parikh-Patel A, et al. First-Line Systemic Treatments for Stage IV Non-Small Cell Lung Cancer in California: Patterns of Care and Outcomes in a Real-World Setting. JNCI Cancer Spectr. 2019;3(3):pkz020. DOI: 10.1093/jncics/pkz020

(29) Norris RP, Dew R, Sharp L, et al. Are there socio-economic inequalities in utilization of predictive biomarker tests and biological and precision therapies for cancer? A systematic review and meta-analysis. BMC Med. 2020;18(1):282-. DOI: 10.1186/s12916-020-01753-0

(30) Marrett E, Kwong WJ, Chow LQ. Factors associated with time to EGFR TKI treatment in patients with non-squamous metastatic non-small-cell lung cancer. Future Oncol. 2022;18(13):1535-44. DOI: 10.2217/fon-2021-1452

(31) McAlister FA, Cram P, Bell CM. Comparing Canadian health care to that in other countries: looking beyond the headlines. CMAJ. 2018;190(8):E207-e8. DOI: 10.1503/cmaj.171527

(32) Institut national d'excellence en santé et en services sociaux (INESSS). Création et caractérisation d'une cohorte québécoise de patients atteints d'un cancer du poumon à l'aide de données clinico-administratives. Québec, Qc: Bibliothèque et Archives nationales du Québec; 2021. p. 152.

(33) Institut national d'excellence en santé et en services sociaux (INESSS). Rapport d'activités scientifiques 2016/2017. Québec, Qc: Bibliothèque et Archives nationales du Québec; 2017.

(34) Canadian Institute for Health Information. Canadian coding standards for version 2022 ICD-10-CA and CCI. CIHI Ottawa, ON; 2022.

(35) Statistics Canada. Canadian classification of diagnostic, therapeutic, and surgical procedures. In: Nosology Reference Centre, editor.: Minister of Supply and Services Canada; 1986.

(36) Statistics Canada. Focus on geography series, 2016 census. Ottawa, Ontario2017.

(37) Gamache P, Hamel D, Blaser C. Material and social deprivation index: A summary. In: populations Bdiedéesd, editor.: Institut national de santé publique du Québec (INSPQ); 2019. p. 16.

(38) Gauthier J, Haggerty J, Lamarche P, et al. Entre adaptabilité et fragilité: les conditions d'accès aux services de santé des communautés rurales et éloignées: Institut national de santé publique du Québec Québec; 2009.

(39) Touati N, Turgeon J. Répartition géographique des médecins de famille: quelles solutions à un problème complexe? Sante Publique (Bucur). 2013;25(4):465-73. DOI: https://doi.org/10.3917/spub.134.0465

(40) Canadian Institutes for Health Information. Population Grouping Methodology [information sheet]. Canadian Institute for Health Information Ottawa; 2020.

(41) Weir S, Steffler M, Li Y, Shaikh S, Wright JG, Kantarevic J. Use of the Population Grouping Methodology of the Canadian Institute for Health Information to predict high-cost health system users in Ontario. CMAJ. 2020;192(32):E907-e12. DOI: 10.1503/cmaj.191297

(42) Korn EL. Censoring distributions as a measure of follow-up in survival analysis. Stat Med. 1986;5(3):255-60. DOI: 10.1002/sim.4780050306

(43) Therneau T. A Package for Survival Analysis in R. 2021. Available from: https://CRAN.R-project.org/package=survival.

(44) Canty AJ. Resampling methods in R: the boot package. The Newsletter of the R Project Volume. 2002;2(3):2-7. DOI:

(45) Papadimitrakopoulou VA, Mok TS, Han JY, et al. Osimertinib versus platinum-pemetrexed for patients with EGFR T790M advanced NSCLC and progression on a prior EGFR-tyrosine kinase inhibitor: AURA3 overall survival analysis. Ann Oncol. 2020;31(11):1536-44. DOI: https://dx.doi.org/10.1016/j.annonc.2020.08.2100

(46) Canadian Partnership Against Cancer. Examining disparities in cancer control: A system performance special focus report. Canadian Partnership Against Cancer; 2014.

(47) Canadian Partnership Against Cancer. Lung cancer and equity: A focus on income and geography. Toronto, ON.: Canadian Partnership Against Cancer; 2020.

(48) Booth CM, Li G, Zhang-Salomons J, Mackillop WJ. The impact of socioeconomic status on stage of cancer at diagnosis and survival. Cancer. 2010;116(17):4160-7. DOI: https://doi.org/10.1002/cncr.25427

(49) Mackillop WJ, Zhang-Salomons J, Groome PA, Paszat L, Holowaty E. Socioeconomic status and cancer survival in Ontario. J Clin Oncol. 1997;15(4):1680-9. DOI: https://doi.org/10.1200/JCO.1997.15.4.1680

(50) Habbous S, Khan Y, Langer DL, 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. DOI: https://doi.org/10.4103/atm.atm_283_20

(51) Finke I, Behrens G, Weisser L, Brenner H, Jansen L. Socioeconomic Differences and Lung Cancer Survival-Systematic Review and Meta-Analysis. Front Oncol. 2018;8:536-. DOI: https://doi.org/10.3389/fonc.2018.00536

(52) Groome PA, Schulze KM, Keller S, Mackillop WJ. Demographic Differences between Cancer Survivors and those who Die Quickly of their Disease. Clin Oncol. 2008;20(8):647-56. DOI: https://doi.org/10.1016/j.clon.2008.05.006

(53) Cheung MC, Earle CC, Fischer HD, et al. Impact of Immigration Status on Cancer Outcomes in Ontario, Canada. J Oncol Pract. 2017;13(7):e602-e12. DOI: https://doi.org/10.1200/JOP.2016.019497

(54) Sorin M, Franco EL, Quesnel-Vallée A. Inter- and intraprovincial inequities in public coverage of cancer drug programs across Canada: a plea for the establishment of a pan-Canadian pharmacare program. Curr Oncol. 2019;26(4):266-9. DOI: 10.3747/co.26.4867

(55) Kasymjanova G, Small D, Cohen V, et al. Lung cancer care trajectory at a Canadian centre: an evaluation of how wait times affect clinical outcomes. Curr Oncol. 2017;24(5):302-9. DOI: 10.3747/co.24.3611

(56) Denault M-H, Labbé C, St-Pierre C, et al. Wait Times and Survival in Lung Cancer Patients across the Province of Quebec, Canada. Current Oncology. 2022;29(5):3187-99. DOI: https://doi.org/10.3390/curroncol29050259

(57) Kasymjanova G, Anwar A, Cohen V, et al. The Impact of COVID-19 on the Diagnosis and Treatment of Lung Cancer at a Canadian Academic Center: A Retrospective Chart Review. Curr Oncol. 2021;28(6):4247-55. DOI: 10.3390/curroncol28060360

(58) Cheema PK, Menjak IB, Winterton-Perks Z, et al. Impact of Reflex EGFR/ALK Testing on Time to Treatment of Patients With Advanced Nonsquamous Non–Small-Cell Lung Cancer. Journal of Oncology Practice. 2017;13(2):e130-e8. DOI: 10.1200/jop.2016.014019

(59) Moroz I, D'Angelo MS. A Blueprint for Optimizing Lung Cancer Care for Better Patient Outcomes. Ottawa: The Conference Board of Canada; 2021.

(60) Agulnik JS, Kasymjanova G, Pepe C, et al. Real-World Pattern of Treatment and Clinical Outcomes of EGFR-Mutant Non-Small Cell Lung Cancer in a Single Academic Centre in Quebec. Current Oncology. 2021;28(6):5179-91. DOI: https://doi.org/10.3390/curroncol28060434

(61) Canada S. Table 13-10-0111-01 Number and rates of new cases of primary cancer, by cancer type, age group and sex. 2021.

(62) Institut national de santé publique du Québec. Portrait de santé du Québec et de ses régions 2006: deuxième rapport national sur l'état de santé de la population du Québec Les statistiques: Institut national de santé publique Québec; 2006.

(63) Government of Canada. Alignment Among Public Formularies in Canada, Part 2: Oncology Medicines. Ottawa, ON.: Patented Medicine Prices Review Board; 2021. p. pp. 1-31.

5.3. Supplementary materials

5.3.1. Appendix A

Intervention codes for medical interventions involving biopsies

Table 5.3.1. Biopsy-related intervention codes in SMOD and MED-ECHO

A. Physician billing codes for fee-for-service renumeration	
00181	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie (unique ou multiple); foie (à l'aiguille, percutanée) (PG-28).
00184	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie (unique ou multiple); ganglion (cervical, axillaire ou inguinal) (PG 28).
00212	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie/cytologie à l'aiguille, par voie transcutanée, sous guidage échoscopique, fluoroscopique ou scanographique; osseuse; os.
00247	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie/cytologie à l'aiguille, par voie transcutanée, sous guidage échoscopique, fluoroscopique ou scanographique; osseuse; vertèbre.
00252	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie/cytologie à l'aiguille, par voie transcutanée, sous guidage échoscopique, fluoroscopique ou scanographique; rétropéritoine (rein, pancréas, ganglions, surrénale).
00273	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie (unique ou multiple); osseuse à l'aiguille (PG-28).
00275	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; RHINOSINUSOLOGIE; Thoracoscopie incluant, le cas échéant, biopsie, section d'adhérences et drainage, thoracique.
00308	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie (unique ou multiple); masse cervicale (à l'aiguille) (PG-28).
00515	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Bronchoscopie rigide incluant la laryngoscopie, la trachéoscopie, la biopsie et l'exérèse de tumeur, le cas échéant; bronchoscopie rigide incluant la biopsie (PG-23).
00524	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Bronchoscopie rigide incluant la laryngoscopie, la trachéoscopie, la biopsie et l'exérèse de tumeur, le cas échéant; médiastinoscopie avec ou sans biopsie.
00592	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES, Ponctions: (incluant injection s'il y a lieu); ganglion, un ou plusieurs.
00753	ACTES DIAGNOSTIQUES ET THÉRAPEUTIQUES; PNEUMOLOGIE; Bronchoscopie incluant la laryngoscopie, la trachéoscopie et la biopsie: localisation bronchoscopique d'un cancer occulte de l'arbre respiratoire, incluant biopsies multiples et aspirations cytologiques multiples, au niveau de toutes les bronches segmentaires.
00782	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie (unique ou multiple); par brossage bronchique (une ou plusieurs bronches) sans usage de bronchoscope ou laryngoscope, incluant l'intubation, l'anesthésie locale et la fluoroscopie.
00797	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie (unique ou multiple); plèvre ou poumon ou les deux: au trépan pneumatique.
02062	MUSCULO-SQUELETTIQUE; SQUELETTE AXIAL; BASSIN; Biopsie osseuse; ouverte.

02066	MUSCULO-SQUELETTIQUE; SQUELETTE AXIAL; THORAX; Biopsie; costale ouverte.
02109	MUSCULO-SQUELETTIQUE; SQUELETTE AXIAL; COLONNE VERTÉBRALE; Biopsie; d'un élément postérieur.
02119	MUSCULO-SQUELETTIQUE; SQUELETTE AXIAL; COLONNE VERTÉBRALE; Biopsie; corps vertébral.
02174	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES SUPÉRIEUR; BRAS (HUMÉRUS); Biopsie osseuse; à l'aiguille.
02175	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES SUPÉRIEUR; BRAS (HUMÉRUS); Biopsie osseuse; ouverte.
02247	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES INFÉRIEURS; PIED; Biopsie; ouverte (PG-28).
02719	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES INFÉRIEURS; FÉMUR; Biopsie; forage et décompression de la tête fémorale.
02796	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES INFÉRIEURS; FÉMUR; Biopsie; à l'aiguille.
02797	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES INFÉRIEURS; FÉMUR; Biopsie; ouverte.
02864	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES INFÉRIEURS; TIBIA ET PÉRONÉ; Biopsie; à l'aiguille.
02865	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES INFÉRIEURS; TIBIA ET PÉRONÉ; Biopsie; ouverte.
02934	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES SUPÉRIEUR; MAIN ET POIGNET; Biopsie; à l'aiguille, main et poignet.
02939	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES SUPÉRIEUR; MAIN ET POIGNET; Biopsie; ouverte, main et poignet (PG-28).
02991	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES SUPÉRIEUR; AVANT-BRAS; Biopsie - radius ou cubitus; à l'aiguille.
02992	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES SUPÉRIEUR; AVANT-BRAS; Biopsie - radius ou cubitus; ouverte.
03011	SYSTÈME RESPIRATOIRE; POUMONS ET PLÈVRE; Incision; Thoracoscopie diagnostique avec ou sans biopsie.
03013	SYSTÈME RESPIRATOIRE; POUMONS ET PLÈVRE; Incision; Thoracoscopie lors d'une autre intervention chirurgicale, supplément – NOTE : Le code 03013 ne peut s'ajouter à un acte comportant la mention « toute voie d'approche » sauf dans le cas d'une thoracoscopie suivie d'une thoracotomie.
03027	SYSTÈME RESPIRATOIRE; MÉDIASTIN; Incision; Médiastinotomie antérieure pour staging avec ouverture de la plèvre incluant la résection costale et le drainage, le cas échéant.
03035	SYSTÈME RESPIRATOIRE; MÉDIASTIN; Incision; Médiastinotomie pour exploration ou drainage : voie cervicale.
03036	SYSTÈME RESPIRATOIRE; MÉDIASTIN; Incision; Médiastinotomie pour exploration ou drainage : voie thoracique.
03120	SYSTÈME RESPIRATOIRE; POUMONS ET PLÈVRE; Incision; Thoracotomie; exploratrice avec biopsie.
03123	SYSTÈME RESPIRATOIRE; POUMONS ET PLÈVRE; Incision; Thoracotomie; exploratrice pour cancer, sans résection, avec ou sans biopsie.

04159	SYSTÈMES LYMPHATIQUE ET HÉMATOPOÏÉTIQUE; Excision; Exérèse de ganglions cervicaux (bénin ou malin) (PG-28).
04161	SYSTÈMES LYMPHATIQUE ET HÉMATOPOÏÉTIQUE; Excision; Excision simple de ganglions lymphatiques pour lésion maligne (PG-28).
07528	SYSTÈMES NERVEUX; CRÂNE & ENCÉPHALE; Lésions expansives tumorales : (incluant les lésions kystiques tumorales); Sus-tentorielles: Biopsie diagnostique par trépanation seulement - (sans stéréotaxie, quel que soit le nombre de biopsies), par le neurochirurgien seulement.
07531	SYSTÈMES NERVEUX; CRÂNE & ENCÉPHALE; Lésions expansives tumorales : (incluant les lésions kystiques tumorales); Sous-tentorielles: Tumeur du tronc cérébral; biopsie seulement.
07643	SYSTÈMES NERVEUX; CRÂNE & ENCÉPHALE; Installation du cadre; ponction ou biopsie (simple ou multiple) de lésion intraparenchymateuse, incluant la technique de localisation.
07690	SYSTÈMES NERVEUX; RACHIS, MOELLE, QUEUE DE CHEVAL; Tumorale; Ponction et/ou biopsie tumorale ou de kyste tumoral ou décompression.
08348	ULTRASONOGRAPHIE; ÉCHOGRAPHIE ARTICULAIRE; Échographie transendoscopique de l'oesophage, de l'estomac, du duodénum ou d'un organe intra-abdominal incluant l'endoscopie gastro-entérologique effectuée avec le scope d'échoendoscopie, maximum de un examen, par jour, par patient.
08349	ULTRASONOGRAPHIE; ÉCHOGRAPHIE ARTICULAIRE; Échographie transendoscopique de l'oesophage, de l'estomac, du duodénum ou d'un organe intra-abdominal incluant l'endoscopie gastro-entérologique effectuée avec le scope d'échoendoscopie; biopsie ou ponction ou injection, unique ou multiple, par voie transoesophagienne, transgastrique ou transduodénale d'une lésion médiastinale ou abdominale, supplément.
08383	ULTRASONOGRAPHIE; ÉCHOGRAPHIE DES VOIES RESPIRATOIRES; échographie endobroinchique.
08384	ULTRASONOGRAPHIE; ÉCHOGRAPHIE DES VOIES RESPIRATOIRES; échographie endobroinchique; avec ponction ganglionnaire ou tumorale transachéale ou transbronchique, supplément.
09367	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Bronchoscopie flexible incluant la laryngoscopie, la trachéoscopie, la biopsie, le lavage broncho-alvéolaire et l'exérèse de tumeur, le cas échéant; Bronchoscopie: avec biopsie pulmonaire transbronchique, supplément.
09368	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Bronchoscopie flexible incluant la laryngoscopie, la trachéoscopie, la biopsie, le lavage broncho-alvéolaire et l'exérèse de tumeur, le cas échéant; Bronchoscopie: avec ponction ganglionnaire transtrachéale et/ou transbronchique, supplément.
09369	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Bronchoscopie incluant la laryngoscopie, la trachéoscopie et la biopsie; bronchoscopie; avec lavage broncho-alvéolaire diagnostique v.g. technique de crystal, supplément.
09418	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Ponctions (incluant injection s'il y a lieu); pleurale, sous guidage, le cas échéant (PG-28).
09464	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie/cytologie à l'aiguille, par voie transcutanée, sous guidage échoscopique, fluoroscopique ou scanographique; thoracique.

09466	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Biopsie/cytologie à l'aiguille, par voie transcutanée, sous guidage échoscopique, fluoroscopique ou scanographique; hépatique.
09502	MUSCULO-SQUELETTIQUE; EXTRÉMITÉS - MEMBRES INFÉRIEURS; PIED; Biopsie; à l'aiguille ou au trocart.
09550	MUSCULO-SQUELETTIQUE; CRÂNE ET FACE; Biopsie; ouverte (unique ou multiple) (PG-28).
20040	PROCÉDÉS DIAGNOSTIQUES ET THÉRAPEUTIQUES; Gastro-entérologie; Endoscopie gastro-entérologique: Oesophagoscopie rigide incluant la biopsie, le cas échéant (PG-23).
B. Canadian classification of health interventions (CCI) codes	
2AA71RW	Biopsie, méninges et duremère du cerveau, approche par trou de trépan avec aspiration à l'aiguille
2AA71SE	Biopsie, méninges et duremère du cerveau, approche par trou de trépan
2AA71SZ	Biopsie, méninges et duremère du cerveau, par craniotomie [volet osseux]
2AC71DA	Biopsie, ventricules cérébraux,approche endoscopique via un trou de trépan (ou la fontanelle)
2AE71DA	Biopsie, thalamus et noyaux gris centraux,approche endoscopique via un trou de trépan
2AE71SE	Biopsie, thalamus et noyaux gris centraux, accès par trou de trépan
2AE71SZ	Biopsie, thalamus et noyaux gris centraux, approche ouverte par craniotomie [volet osseux]
2AF71GR	Biopsie, région pituitaire, approche transluminale percutanée
2AF71QS	Biopsie, région pituitaire, approche transsphénoïdale ouverte [transethmoïdale]
2AG71SE	Biopsie, glande pinéale, approche par trou de trépan
2AJ71SE	Biopsie, cervelet, approche par trou de trépan (pour accéder au cervelet)
2AJ71SZ	Biopsie, cervelet, approche ouverte par craniotomie ou craniectomie [volet osseux]
2AN71SE	Biopsie, cerveau, approche par trou de trépan
2AN71SZ	Biopsie, cerveau, approche ouverte par craniotomie ou craniectomie [volet osseux]
2AP71SZ	Biopsie, tronc cérébral, approche ouverte par craniotomie ou craniectomie [volet osseux]
2AW71DA	Biopsie, moelle épinière, approche endoscopique (laparoscopie)
2AW71HA	Biopsie, moelle épinière,approche percutanée (à l'aiguille)
2AW71LA	Biopsie, moelle épinière,approche ouverte
2AX71DA	Biopsie, canal rachidien et méninges, approche endoscopique [laparoscopie]
2AX71HA	Biopsie, canal rachidien et méninges, approche percutanée
2AX71LA	Biopsie, canal rachidien et méninges, approche ouverte
2BG71DC	Biopsie, plexus brachial, approche endoscopique (thoracoscopie) par abord antérieur
2BG71LL	Biopsie, plexus brachial, approche ouverte antérieure
2EA71HA	Biopsie, crâne, approche percutanée (à l'aiguille)
2EA71LA	Biopsie, crâne, approche ouverte
2EB71LA	Biopsie, apophyse zygomatique,approche ouverte
2ED71LA	Biopsie, maxillaire,approche ouverte
2EE71LA	Biopsie, mandibule,approche ouverte
2GM70LA	Inspection, bronches,approche ouverte
2GM71BA	Biopsie, bronches,approche endoscopique par voie naturelle
2GM71BP	Biopsie, bronches approche endoscopique par voie naturelle, avec aspiration à l' aiguille

2GM71BR	Biopsie, bronches, approche endoscopique par voie naturelle, brossage ou lavage
2GM71LA	Biopsie, bronches,approche ouverte
2GT70LA	Inspection, poumon NCA, approche ouverte
2GT71BA	Biopsie, poumon NCA, approche endoscopique par voie naturelle
2GT71BP	Biopsie, poumon NCA, approche endoscopique par voie naturelle et biopsie par aspiration
2GT71DA	Biopsie, poumon NCA, approche endoscopique [chirurgie thoracique vidéo assistée]
2GT71HA	Biopsie, poumon NCA, approche percutanée (à l'aiguille)
2GT71LA	Biopsie, poumon NCA, approche ouverte
2GW71DA	Biopsie, médiastin, approche endoscopique [chirurgie thoracique vidéo assistée]
2GW71HA	Biopsie, médiastin,approche percutanée (à l'aiguille)
2GW71LA	Biopsie, médiastin,approche ouverte
2GY70DA	Inspection, cavité thoracique NCA, approche endoscopique [chirurgie thoracique vidéo assistée]
2GY70LA	Inspection, cavité thoracique NCA, approche ouverte
2MA71HA	Biopsie, ganglion(s) lymphatique(s), région de la tête, approche percutanée (à l'aiguille)
2MA71LA	Biopsie, ganglion(s) lymphatique(s), région de la tête, approche ouverte
2MB71HA	Biopsie, ganglions lymphatiques de l'artère cervicale profonde,approche percutanée (à l'aiguille)
2MB71LA	Biopsie, ganglions lymphatiques de l'artère cervicale profonde,approche ouverte
2MC71HA	Biopsie, ganglion(s) lymphatique(s), cervical(aux), approche percutanée (à l'aiguille)
2MC71LA	Biopsie, ganglion(s) lymphatique(s), cervical(aux), approche ouverte
2MD71HA	Biopsie, ganglion(s) lymphatique(s), axillaire(s), approche percutanée (à l'aiguille)
2MD71LA	Biopsie, ganglion(s) lymphatique(s), axillaire(s), approche ouverte
2ME71BP	Biopsie, ganglion(s) lymphatique(s), médiastinal(aux), approche endoscopique par voie naturelle avec aspiration à l'aiguille
2ME71DA	Biopsie, ganglion(s) lymphatique(s), médiastinal(aux), approche endoscopique
2ME71HA	Biopsie, ganglion(s) lymphatique(s), médiastinal(aux), approche percutanée (à l'aiguille)
2ME71LA	Biopsie, ganglion(s) lymphatique(s), médiastinal(aux), approche ouverte
2MF71DA	Biopsie, ganglion(s) lymphatique(s), intrathoracique(s) NCA, approche endoscopique
2MF71HA	Biopsie, ganglion(s) lymphatique(s), intrathoracique(s) NCA, approche percutanée (à l'aiguille)
2MF71LA	Biopsie, ganglion(s) lymphatique(s), intrathoracique(s) NCA, approche ouverte
2MG71DA	Biopsie, ganglion(s) lymphatique(s), intraabdominal(aux), approche endoscopique
2MG71HA	Biopsie, ganglion(s) lymphatique(s), intraabdominal(aux), approche percutanée (à l'aiguille)
2MG71LA	Biopsie, ganglion(s) lymphatique(s), intraabdominal(aux), approche ouverte
2MH71DA	Biopsie, ganglion(s) lymphatique(s), pelvien(s), approche endoscopique
2MH71HA	Biopsie, ganglion(s) lymphatique(s), pelvien(s), approche percutanée (à l'aiguille)
2MH71LA	Biopsie, ganglion(s) lymphatique(s), pelvien(s), approche ouverte
2MJ71DA	Biopsie, ganglion(s) lymphatique(s), inguinal(aux), approche endoscopique
2MJ71HA	Biopsie, ganglion(s) lymphatique(s), inguinal(aux), approche percutanée (à l'aiguille)
2MJ71LA	Biopsie, ganglion(s) lymphatique(s), inguinal(aux), approche ouverte
2MK71HA	Biopsie, ganglion(s) lymphatique(s), membre(s) NCA, approche percutanée (à l'aiguille)
2MK71LA	Biopsie, ganglion(s) lymphatique(s), membre(s) NCA, approche ouverte

2MZ71HA	Biopsie, système lymphatique,approche percutanée (à l'aiguille)
2MZ71LA	Biopsie, système lymphatique,approche ouverte
2OA71DA	Biopsie, foie, approche endoscopique (laparoscopie)
2OA71GR	Biopsie, foie, approche veineuse transluminale percutanée
2OA71HA	Biopsie, foie,approche percutanée (à l'aiguille)
2OA71LA	Biopsie, foie,approche ouverte
2PB71DA	Biopsie, glande surrénale,approche endoscopique
2PB71HA	Biopsie, glande surrénale,approche percutanée (à l'aiguille)
2PB71LA	Biopsie, glande surrénale,approche ouverte
2SA71HA	Biopsie, atlas et axe,approche percutanée (à l'aiguille)
2SC71HA	Biopsie, vertèbres,approche percutanée (à l'aiguille)
2SC71LL	Biopsie, vertèbres, approche antérieure ouverte
2SC71PF	Biopsie, vertèbres, approche postérieure ouverte (postérolatérale)
2SE71HA	Biopsie, disque intervertébral, approche percutanée (à l'aiguille)
2SF71HA	Biopsie, sacrum et coccyx, approche percutanée (à l'aiguille)
2SF71LA	Biopsie, sacrum et coccyx, approche ouverte
2SK71HA	Biopsie, sternum,approche percutanée (à l'aiguille)
2SK71LA	Biopsie, sternum,approche ouverte
2SL71HA	Biopsie, côtes,approche percutanée (à l'aiguille)
2SL71LA	Biopsie, côtes,approche ouverte
2SM71HA	Biopsie, clavicule,approche percutanée (à l'aiguille)
2SM71LA	Biopsie, clavicule,approche ouverte
2SN71HA	Biopsie, omoplate,approche percutanée (à l'aiguille)
2SQ71HA	Biopsie, bassin,approche percutanée (à l'aiguille)
2SQ71LA	Biopsie, bassin,approche ouverte
2SW71HA	Biopsie, pubis,approche percutanée (à l'aiguille)
2SW71LA	Biopsie, pubis,approche ouverte
2TK71HA	Biopsie, humérus,approche percutanée (à l'aiguille)
2TK71LA	Biopsie, humérus,approche ouverte
2TV71LA	Biopsie, radius et cubitus,approche ouverte
2UJ71LA	Biopsie, phalanges de la main, approche ouverte
2VC71HA	Biopsie, fémur,approche percutanée (à l'aiguille)
2VC71LA	Biopsie, fémur,approche ouverte
2VP71HA	Biopsie, rotule,approche percutanée (à l'aiguille)
2VQ71HA	Biopsie, tibia et péroné,approche percutanée (à l'aiguille)
2VQ71LA	Biopsie, tibia et péroné,approche ouverte
2WL71LA	Biopsie, phalange du pied,approche ouverte
2WS71HA	Biopsie, métatarses,approche percutanée (à l'aiguille)
3GY30HJ	Ultrason, cavité thoracique NCA, par voie transoesophagienne
C. Canadian classification of diagnostic, therapeutic, and surgical procedures (CCP) codes	
1682	Biopsie de la moelle et des méninges
2081	Biopsie percutanée (à l'aiguille) de la glande surrénale
2082	Autre biopsie de la glande surrénale

4581	Biopsie bronche / Biopsie des bronches par bronchoscopie
4582	Autre biopsie des bronches
4583	Biopsie percutanée (à l'aiguille) du poumon
4602	Thoracotomie d'exploration
4619	Médiastinotomie / Incision de médiastin
4681	Thoracoscopie / Thoracoscopie transpleurale
4682	Médiastinoscopie
4683	Biopsie de la paroi thoracique
4685	Biopsie percutanée (à l'aiguille) de médiastin
4686	Biopsie médiastine
4691	Thoracocentèse / Ponction pleurale
5281	Biopsie d'une formation lymphatique
6281	Biopsie percutanée du foie
6282	Autre biopsie due foie
8881	Biopsie d'os de la face
8990	Biopsie osseuse omoplate, calvicule, thorax, côtes, sternum
8991	Biopsie de l'os, humérus
8992	Biopsie osseuse radius, cubitus (avant-bras)
8994	Biopsie osseuse fémur
8996	Biopsie osseuse tibia, péroné (cheville)
8998	Biopsie osseuse
8999	Biopsie osseuse siège autre ou non précisé
14811	Biopsie fermée [percutanée] [à l'aiguille] des méninges cérébrales
14821	Biopsie fermée [percutanée] [à l'aiguille] du cerveau
14822	Biopsie ouverte cérébroméningée

5.3.2. Appendix B

Details on exposures of interest

Table 5.3.2. Exposures of Interest

Name	Categories
Median income after tax[1]	Low: 0$ - 28,222$ Medium: 28,223$ - 31,653$ High: ≥31,654$
Percent with low education[1,2]	Low: 0 - 28.69 % Medium: 28.7 - 34.99 % High: ≥ 35.00 %
Material deprivation index	Low: Tercile 1 (least deprived) Medium: Tercile 2 High: Tercile 3 (most deprived)
Percent immigrant[1]	Low: 0 - 4.09 % Medium: 4.10 -10.29 % High: ≥ 10.30 %
Census-based region[3]	Urban (CMA): Montreal Urban (CMA): Others Suburban (CA) Rural (Non-CMA/CA)
Health center-based region	University Peripheral Intermediary Remote or Northern

Abbreviations: CMA: Census Metropolitan Area; CA: Census Agglomeration. [1] The three categories were created using quintiles: quintiles 1 and 2 combined as the "Low" group, quintile 3 as the "Medium" group, and quintiles 4 and 5 combined as the "High" group; [2] Low education is defined as no secondary diploma or an equivalent attestation for individuals between 15-24 or ≥65 years, and no certificate or no post-secondary education for individuals between 25-64 years; [3] The "CMA: Others" category includes the following cities: Quebec, Trois-Rivières, Gatineau, Saguenay, and Sherbrooke.

5.3.3. Appendix C

Balance in distributions of age and sex before and after inverse weighting with propensity scores

Table 5.3.3. Distributions of age and sex in cohort 1 before and after inverse propensity weighting

Propensity Score Weighting	Population	Age (Years) n (%)				Sex n (%)	
		< 60 years	60-64 years	65-70 years	≥ 70 years	Male	Female
Unweighted	**Overall (N=457)**	58 (12.7)	48 (10.5)	102 (22.3)	249 (54.5)	144 (31.5)	313 (68.5)
Unweighted	Income[1] – Low	28 (14.0)	19 (9.5)	40 (20.0)	113 (56.5)	67 (33.5)	133 (66.5)
	Income[1] - Medium	13 (13.1)	11 (11.1)	21 (21.2)	54 (54.5)	31 (31.3)	68 (68.7)
	Income[1] – High	17 (11.3)	17 (11.3)	39 (25.8)	78 (51.7)	45 (29.8)	106 (70.2)
	Low Education[2] – Low	23 (11.1)	19 (9.2)	50 (24.3)	114 (55.3)	58 (28.2)	148 (71.8)
	Low Education[2] - Medium	10 (13.0)	10 (13.0)	17 (22.1)	40 (51.9)	30 (39.0)	47 (61.0)
	Low Education[2] - High	25 (14.6)	18 (10.5)	35 (20.5)	93 (54.4)	55 (32.2)	116 (67.8)
	Material Deprivation[3] - Low	18 (12.2)	13 (8.8)	37 (25.2)	79 (53.7)	37 (25.2)	110 (74.8)
	Material Deprivation[3] - Medium	15 (11.9)	18 (14.3)	29 (23.0)	64 (50.8)	48 (38.0)	78 (61.9)
	Material Deprivation[3] - High	22 (14.5)	15 (9.9)	32 (21.1)	83 (54.6)	50 (32.9)	102 (67.1)
	Immigration[4] – Low	15 (9.9)	14 (9.2)	45 (29.6)	78 (51.3)	53 (34.9)	99 (65.1)
	Immigration[4] - Medium	11 (15.7)	6 (8.6)	15 (21.4)	38 (54.3)	19 (27.1)	51 (72.9)
	Immigration[4] - High	32 (13.8)	27 (11.6)	42 (18.1)	131 (56.5)	71 (30.6)	161 (69.4)
	Census-Based Region - Urban: Montreal	36 (14.1)	33 (13.0)	50 (19.7)	135 (53.1)	77 (30.3)	177 (69.7)
	Census-Based Region - Urban: Others	11 (13.3)	4 (4.8)	19 (22.9)	49 (59.0)	25 (30.1)	58 (69.9)
	Census-Based Region - Suburban	6 (14.6)	2 (4.9)	10 (24.4)	23 (56.1)	14 (34.1)	27 (65.9)
	Census-Based Region - Rural	5 (6.5)	8 (10.4)	23 (29.9)	41 (53.2)	28 (36.4)	49 (63.6)
	Health Center-Based Region - University	31 (13.3)	24 (10.3)	40 (17.2)	138 (59.2)	72 (30.9)	161 (69.0)
	Health Center-Based Region - Peripheral	21 (14.8)	19 (13.4)	43 (30.3)	59 (41.5)	41 (28.9)	101 (71.1)
	Health Center-Based Region – Intermediary, Remote or Northern	6 (7.5)	4 (5.0)	19 (23.8)	51 (63.8)	31 (38.8)	49 (61.3)

Weighted	Income[1] - Low	25.9 (13.0)	20.6 (10.3)	44.4 (22.2)	109.0 (54.5)	64,7 (32.4)	135.2 (67.6)
	Income[1] - Medium	12.8 (12.9)	10.4 (10.5)	22.0 (22.2)	53.8 (54.5)	31.7 (32.0)	67.3 (68.0)
	Income[1] - High	19.3 (12.8)	16.0 (10.6)	32.9 (21.8)	82.7 (54.8)	46.9 (31.0)	104.1 (69.0)
	Low Education[2] - Low	26.5 (12.9)	20.8 (10.1)	46.4 (22.5)	112.2 (54.5)	65.0 (31.6)	140.8 (68.4)
	Low Education[2] - Medium	10.0 (13.0)	8.1 (10.5)	17.5 (22.8)	41.3 (53.7)	24.4 (31.7)	52.5 (68.3)
	Low Education[2] - High	21.8 (12.8)	17.7 (10.4)	38.5 (22.5)	93.0 (54.4)	54.9 (32.1)	116.1 (67.9)
	Material Deprivation[3] - Low	18.3 (12.7)	16.5 (11.4)	35.1 (24.2)	75.0 (51.8)	42.1 (29.0)	102.9 (71.0)
	Material Deprivation[3] - Medium	16.6 (12.9)	16.0 (12.4)	25.1 (19.5)	71.0 (55.2)	49.4 (38.4)	79.3 (61.6)
	Material Deprivation[3] - High	20.6 (13.6)	13.7 (9.1)	37.0 (24.4)	80.2 (52.9)	44.8 (29.6)	106.7 (70.4)
	Immigration[4] - Low	18.8 (12.4)	15.9 (10.5)	34.2 (22.6)	82.8 (54.6)	48.9 (32.2)	102.8 (67.8)
	Immigration[4] - Medium	8.7 (12.4)	7.0 (10.0)	16.3 (23.2)	38.2 (54.4)	22.3 (31.8)	47.8 (68.2)
	Immigration[4] - High	29.6 (12.4)	24.1 (10.4)	51.4 (22.2)	126.8 (54.7)	73.3 (31.6)	158.6 (68.4)
	Census-Based Region - Urban: Montreal	32.4 (12.7)	26.3 (10.4)	57.0 (22.4)	138.4 (54.5)	80.2 (31.6)	173.8 (68.4)
	Census-Based Region - Urban: Others	10.6 (12.8)	8.4 (10.1)	18.6 (22.4)	45.3 (54.7)	25.3 (30.5)	57.6 (69.5)
	Census-Based Region - Suburban	5.2 (12.7)	4.3 (10.5)	9.2 (22.5)	22.3 (54.3)	12.5 (30.6)	28.5 (69.4)
	Census-Based Region - Rural	9.7 (12.6)	8.1 (10.5)	17.3 (22.5)	41.8 (54.4)	24.8 (32.2)	52.1 (67.8)
	Health Center-Based Region - University	29.7 (12.7)	24.1 (10.3)	52.3 (22.4)	126.9 (54.5)	74.7 (32.1)	158.3 (67.9)
	Health Center-Based Region - Peripheral	18.1 (12.7)	14.8 (10.4)	31.8 (22.4)	77.3 (54.4)	44.5 (31.3)	97.4 (68.7)
	Health Center-Based Region – Intermediary, Remote or Northern	8.2 (10.6)	7.0 (9.1)	18.6 (24.2)	42.9 (56.0)	23.7 (31.0)	52.9 (69.0)

[1]Low: 0 - 28,222$; Medium: 28,223$ - 31,653$; High: ≥ 31,654$; [2]Low: 0 - 28.69 %; Medium: 28.7 - 34.99 %; High: ≥ 35.0 %; [3]Low: Tercile 1; Medium: Tercile 2; High: Tercile 3; [4]Low: 0 - 4.09 %; Medium: 4.10 – 10.29 %; High: ≥ 10.30 %.

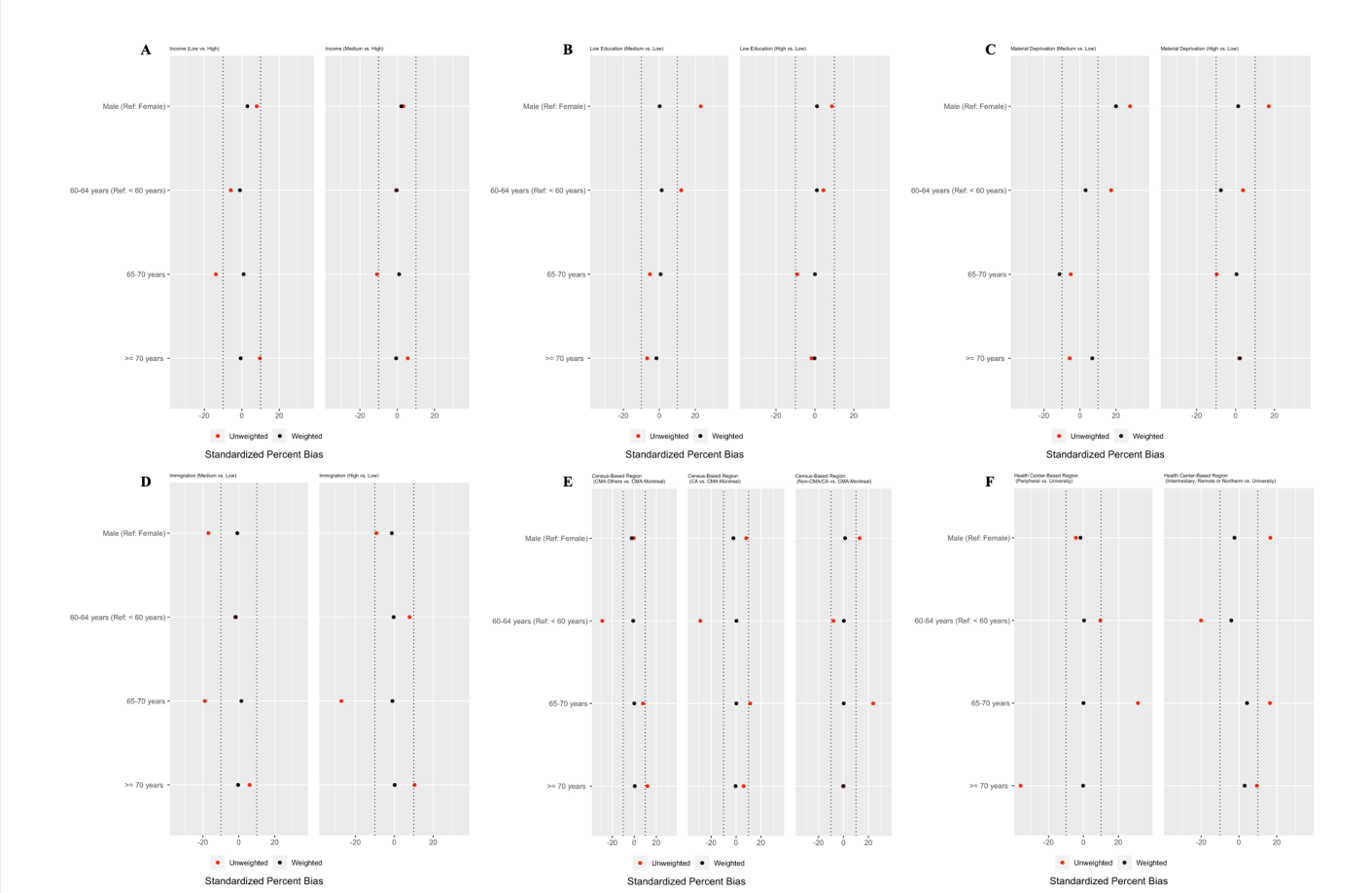


Figure 5.3.1. Balance in distributions of age and sex in cohort 1 before and after propensity score weighting
A: Income; B: Low Education; C: Material Deprivation; D: Immigration; E: Census-Based Region; F: Health Center-Based Region. N=457

Standardized Percent Bias = $100 * \left(\frac{\hat{p}_{treatment} - \hat{p}_{control}}{\sqrt{\frac{\hat{p}_{treatment}(1-\hat{p}_{treatment}) + \hat{p}_{control}(1-\hat{p}_{control})}{2}}} \right)$, where $\hat{p}_{treatment}$ and $\hat{p}_{control}$ represent probabilities of the covariate in the comparison and reference groups, respectively, of a categorical treatment variable.

5.3.4. Appendix D

Secondary analysis on the likelihood of receiving osimertinib in cohort 2

Table 5.3.4. Receiving osimertinib as a 2nd EGFR-TKI after receiving gefitinib

	Main Analysis[1] [Cohort 1, N = 457]		Secondary Analysis[1] [Cohort 2, N = 195]	
	Percentage (95% CI)	**Percentage Ratio (95% CI)**	**Percentage (95% CI)**	**Percentage Ratio (95% CI)**
Income				
Low: 0 - 28,222$	17.8 (12.5 - 22.6)	0.83 (0.53 - 1.29)	35.6 (25.8 - 47.8)	0.84 (0.56 - 1.30)
Medium: 28,223$ - 31,653$	22.2 (14.4 - 30.4)	1.03 (0.62 - 1.68)	48.8 (33.3 - 63.7)	1.15 (0.72 - 1.80)
High: ≥ 31,654$	21.5 (14.7 - 28.8)	REF	42.3 (30.3 - 53.5)	REF
Low Education				
Low: 0 - 28.69 %	24.3 (18.6 - 29.5)	REF	47.0 (37.7 - 56.5)	REF
Medium: 28.7 - 34.99 %	15.8 (7.6 - 24.2)	0.65 (0.31 - 1.14)	36.8 (21.0 - 56.7)	0.78 (0.43 - 1.18)
High: ≥ 35.00 %	16.9 (11.1 - 22.8)	0.69 (0.43 - 1.07)	34.8 (23.0 - 46.0)	0.74 (0.48 - 1.08)
Material Deprivation Index				
Low (Least Deprived)	21.7 (15.6 - 29.1)	REF	43.1 (30.6- 54.6)	REF
Medium	24.2 (16.8 - 32.9)	1.12 (0.69 - 1.82)	45.1 (32.2 - 57.6)	1.05 (0.69 - 1.63)
High (Most Deprived)	16.3 (10.3 - 22.2)	0.75 (0.43 - 1.22)	32.8 (20.6 - 45.0)	0.76 (0.46 - 1.15)
Immigration				
Low: 0 - 4.09 %	12.7 (7.4 - 18.1)	REF	27.5 (17.0 - 38.4)	REF
Medium: 4.10 - 10.29 %	19.8 (11.0 - 28.8)	1.56 (0.75 - 3.03)	36.4 (20.0 - 54.1)	1.32 (0.67 - 2.51)
High: ≥ 10.30 %	24.8 (19.4 - 29.9)	1.95 (1.26 - 3.36)	50.9 (40.8 - 61.5)	1.85 (1.22- 3.13)
Census-Based Region				
Urban: Montreal	24.1 (18.5 - 29.6)	REF	47.1 (37.8 - 57.2)	REF
Urban: Others	9.4 (3.8 - 16.1)	0.39 (0.16 - 0.71)	26.6 (10.8 - 45.9)	0.56 (0.23 - 0.98)
Suburban	28.9 (14.7 - 44.5)	1.20 (0.60 - 2.05)	54.7 (33.7 - 77.7)	1.16 (0.70 - 1.75)
Rural	13.3 (5.9 - 22.1)	0.55 (0.25 - 0.98)	26.0 (10.9 - 41.7)	0.55 (0.24 - 0.90)
Health Center-Based Region				
University	19.8 (14.4 - 25.4)	REF	43.8 (34.1 - 54.6)	REF
Peripheral	22.7 (15.3 - 29.0)	1.15 (0.74 - 1.71)	44.3 (31.2 - 56.2)	1.01 (0.66 - 1.44)
Intermediary, Remote or Northern	16.1 (8.1 - 24.9)	0.81 (0.42 - 1.38)	28.8 (16.8 - 44.3)	0.66 (0.36 - 1.06)

[1]Analyses accounted for age (<60, 60-64, 65-70, ≥70 years) and sex. Abbreviations: SES: Socioeconomic status; SDS: Sociodemographic status.

Chapter 6. Manuscript 3

6.1. Preface

This manuscript builds on earlier studies by shifting from examining the impact of individual social identities on lung cancer survival to adopting a comprehensive, intersectional approach. Secondary analyses in Manuscript 1 suggested that the effects of low income and immigrant status may be more pronounced in women. Manuscript 2 demonstrated that individuals from areas with the poorest employment rates, income, and education levels have significantly lower survival rates relative to those in the least deprived areas. These findings motivated a deeper exploration of how intersecting social identities—such as age, sex, income, and immigration status—shape lung cancer survival, alongside contextual factors.

This study was conducted using a robust methodology for quantitative intersectional analysis, leveraging a large-scale administrative database linkage project recently made accessible through Statistics Canada's RDCs. I am among the first users of this dataset, which parallels the linkage project in Manuscript 1 but includes all individuals in the CCR rather than only survey participants. Moreover, the dataset integrates the IMDB, enabling the identification of immigrants diagnosed with cancer in Canada.

The findings from this study reveal inequalities in lung cancer survival driven by intersecting social identities and underscore key populations for advancing research and, subsequently, policy aimed at reducing inequities in lung cancer outcomes.

6.2. Age, sex, income and immigration status and lung cancer survival in Canada: a population-level intersectional multilevel analysis

Samia Qureshi[1*], Arijit Nandi[1,2], and Erin Strumpf[1,3]

[1]Department of Epidemiology, Biostatistics and Occupational Health (EBOH), McGill University, H3A 1G1, Montreal, Quebec, Canada

[2] Department of Equity, Ethics, and Policy, McGill University, H3A 1G1, Montreal, Quebec, Canada

[3] Department of Economics, McGill University, H3A 1G1, Montreal, Quebec, Canada

[*]**Corresponding Author:** Samia Qureshi

Address for Correspondence: samia.qureshi@mail.mcgill.ca

Key Words: Lung cancer, survival, health inequalities, socioeconomic factors, sociodemographic factors, income, immigrant status, intersectionality, multilevel analysis, I-MAIHDA

Word Count: Abstract: 302**;** Main text: 6,064

Abstract

Inequalities in lung cancer outcomes persist across diverse social identities in Canada. While research on individual social identities is expanding, the impact of intersecting social identities on lung cancer outcomes remains largely unexplored. This study addresses this gap by analyzing lung cancer survival at the intersections of four social identities while also considering place of residence and temporal trends.

Using the Canadian Cancer Registry linked to other administrative databases, we aimed to identify all individuals diagnosed with lung cancer from 2002 to 2010. We applied an intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy (I-MAIHDA) to assess 1-year survival probabilities, including intersectional effects, across 72 strata based on age, sex, income, and immigration status.

Among 183,350 individuals diagnosed with lung cancer, 1-year survival probabilities ranged from 23.3% to 66.9% across intersectional strata. Low-income, long-term resident men generally showed the lowest survival probabilities compared to other strata, conditional on age. In contrast, immigrant women exhibited some of the highest survival probabilities. The model's discriminatory accuracy (VPC=5.7%) was low, with 98.2% of between-strata variance explained by the additive effects of each social identity. A closer inspection of stratum-specific estimates revealed that low-income, long-term resident men under 60 and low-income immigrant women aged 80 or older had survival probabilities that were 2.8 percentage points lower than expected based solely on additive effects (95% CI: -5.0 to -0.5 and -5.40 to -0.28, respectively). Most intersectional effects diminished after incorporating stage at diagnosis into the models.

Although the evidence for overall intersectionality for 1-year survival was modest, findings from this study do demonstrate the applicability of intersectional analyses to uncover select key populations experiencing unique social disadvantages in lung cancer survival. Identifying these populations is critical for advancing research on cancer inequalities, which can subsequently inform targeted interventions and policies to mitigate disparities and improve outcomes.

Introduction

Lung cancer remains the leading cause of cancer deaths in Canada, with survival outcomes that vary according to socioeconomic or sociodemographic status.[1-7] However, social identities do not operate in isolation within individuals; instead, they may intersect or overlap to shape survival experiences.

According to the World Health Organization, social determinants of health - such as age, gender, income, and immigrant status - are "conditions in which people are born, grow, live, work, and age."[8-11] These social positions are influenced by broader political, social, and cultural forces and norms that can place individuals at increased risk of poor health - a concept referred to as structural vulnerability.[8, 12] Consequently, inequities in health often stem from structural violence: the harmful arrangement of power, money, and resources, including access to healthcare.[8, 12, 13] Furthermore, intersectionality theory posits that an individual's social location is shaped by multiple, intertwined social dimensions that can create unique forms of oppression or privilege. It is a critical theory rooted in activism and provides a framework for examining and addressing health inequities through a social justice lens.[14]

Despite Canada's universal healthcare access at the point of care, age, sex, income, and immigration status systematically impact cancer care access and outcomes.[15, 16] Older cancer patients are often undertreated.[17, 18] This is also seen with older lung cancer patients,[19-22] even though treatment at older ages does not imply worse outcomes.[23] Age-based discrimination in treatment may be driven by physician preferences or biases, a perceived lack of treatment benefits, and treatment rationing.[15, 17, 21, 22, 24, 25] Women are more likely to receive palliative care,[17] and be diagnosed earlier with lung cancer,[2, 26] potentially due to greater engagement with the healthcare system, while men have lower rates of realized healthcare access.[26-29] Individuals of low socioeconomic status often receive suboptimal preventive and therapeutic services along the cancer care continuum,[15-17] which may stem from out-of-pocket costs, limited supportive and coordinated care, and dismissive or judgmental attitudes, particularly toward patients with a history of substance abuse.[30, 31] Recent immigrants underutilize preventive services and lung cancer treatment, potentially due to language or cultural barriers and a lack of contact with or knowledge of the healthcare system.[20, 32-34]

Seniors who live alone are more likely to be low-income,[35] and studies indicate that they are also more likely to forego cancer treatments.[36, 37] Studies also suggest that age disparities in lung cancer survival differ by sex.[38] Also, women who are low-income and recent immigrants have the lowest rates of cervical cancer screening relative to all other women.[39] Similarly, South Asian immigrant women have the poorest breast and cervical cancer screening rates, which are a result of the complex interplay between multiple social identities (e.g., immigration status, age, gender, race, class, and religion).[40] These are just some examples of how the intersection of multiple structurally vulnerable social identities can lead to unique cancer experiences. Although the need to consider multiple social identities through an intersectional lens is widely recognized in research on cancer care inequities,[16, 31, 41-44] there has been limited progress in applying this concept to quantitative research, including studies on inequalities in lung cancer survival.[5, 7]

To address health inequities effectively, it is essential to consider outcome heterogeneity between and within unique social identities. For example, low income may decrease individuals' chances of survival relative to high income, but the chances of survival among low-income individuals may not be homogeneous due to immigrant status. In this study, we applied a descriptive quantitative intersectional analysis with an intercategorical approach to assess survival across strata defined by four social identities: age, sex, income, and immigration status, while taking into consideration the the context of place and time. Our analyses centred on 1) assessing the discriminatory power of these strata for 1-year survival probabilities, 2) examining the variation in predicted 1-year survival probabilities across strata, and 3) estimating interactions (i.e., intersectional or multiplicative effects) of each strata by assessing deviations in survival probabilities from expectations based solely on the additive effects of individual social identities. This approach aims to highlight segments of the lung cancer population experiencing distinct survival advantages or disadvantages to advance research needed to promote equity in cancer care.

Material and Methods

Study sample

For this retrospective cohort analysis, we included all individuals diagnosed with primary lung cancer in the Canadian Cancer Registry (CCR) between 2002 and 2013 (see details in Appendix A). Individuals diagnosed in Quebec were omitted, as the CCR files used in this study did not contain incidence data from Quebec beyond 2010, which covers part of our study period. We excluded all individuals diagnosed with stage 0, those missing information on date of birth and sex, and those missing data on income, place of residence, date of diagnosis, or date of death after imputation. The resulting sample was used for both primary and secondary analyses, with a subset of individuals diagnosed between 2010 and 2013 used for additional secondary analyses.

Data sources

The CCR incidence Tabulation Master File (CCR TMF: 1992-2015) was linked to the T1 Family File, which includes income tax return data (T1FF: 1992-2015), the Longitudinal Immigration Database (IMDB: 1980-2015), the Canadian Vital Statistics Death Database (1992-2014), death information from the T1 Personal Master File (i.e. income tax return data, T1PMF: 1992-2014), the Discharge Abstract Database (DAD: 1994/95-2015/16), and the National Ambulatory Care Reporting System (NACRS: 2002/03-2015/16). The Postal Code Conversion File Plus (PCCF+) program in SAS was used for geographic coding of postal codes in the CCR and the T1FF (see details in Appendix A).[45]

Main exposures

Information on age and sex was obtained directly from the CCR at the time of diagnosis. Sex was dichotomous (1: male, 2: female), and age was derived in years based on the date of birth and date of diagnosis and divided into six categories (1: <60 years, 2: 60-64 years, 3: 65-69 years, 4: 70-74 years, 5: 75-79 years, and 6: ≥80 years). If only the day was missing for the date of birth or diagnosis, it was imputed as the 15th of the month. We used the adjusted after-tax family income from T1FF, corresponding to the taxation year two years prior to the diagnosis year. Single imputation was used for missing income when possible (see details in Appendix A). Family income was ranked into terciles based on population-level references for each calendar year, province, and place of residence (i.e., small urban cities, large urban cities, rural areas, and territories), and categorized as high income (1), medium income (2), or low income (3). All

individuals identified in the IMDB were classified as immigrants, while those not in the database were considered long-term residents, including Canadian-born individuals or those who landed in Canada before 1980 (1: long-term residents, 2: immigrants) (see details in Appendix A).

All categories of age, sex, income, and immigration status were combined to create 72 mutually exclusive intersectional strata. For simplicity, strata were identified with 4-digit codes: the first digit for age, the second for sex, the third for family-income tercile, and the fourth for immigration status. For example, the code '1231' represents less than 60-year-old women with low income who are long-term residents of Canada.

Outcome

Information on death was derived from the CVSD, T1PMF, or CCR (see details in Appendix A). All individuals included in our study were followed for survival until December 31, 2014, ensuring at least a 1-year follow-up for everyone. Those who died within this 1-year period were coded as dead (0), and those who survived for one year or beyond were coded as alive (1). If only the day was missing from the date of death, it was imputed as the 15th of the month. We did not examine death at longer follow-up periods, like 2 years, since this would have required excluding individuals diagnosed in 2013 to ensure complete follow-up, thereby reducing the power of our analyses.

Covariates

Covariates included in the main or secondary analyses were time period, place of residence, histology, stage, and comorbidities. Calendar year was divided into three categories (1: 2002-2005, 2006-2010, and 2011-2013) to account for improvements in survival over time[46, 47] and the introduction of targeted therapies.[48] Individuals' postal codes at the time of diagnosis were used to identify place of residence in terms of province and in terms of Statistical Area Classification (SAC) types: a grouping of census subdivisions, which are standard geographic areas used to disseminate census data (see details in Appendix A). The SAC types were grouped into 1) large urban cities (census metropolitan areas), 2) small urban cities (census agglomerations), and 3) rural areas (rural areas with none to strong metropolitan influence or territories).

We derived histology of each tumour by categorizing histological codes based on the International Classification of Diseases for Oncology, third edition (ICD-O-3), into seven groups: 1) squamous

cell carcinoma, 2) adenocarcinoma, 3) small cell carcinoma, 4) large cell carcinoma, 5) non-small cell lung cancer, not otherwise specified, 6) sarcomas and other specified neoplasms, and 7) unspecified (see details in Appendix A). Cancer stage has been systematically collected in the CCR since 2010, and we used cancer stage data established through the Collaborative Stage data collection system, which is based on tumour–node–metastasis definitions in the seventh edition of the American Joint Committee on Cancer Staging Manual.[49] Stage I and II cancers were classified as early stage, while stages III and IV were classified as locally advanced or metastatic.

Diagnostic codes based on the International Classification of Diseases, 10th Revision, with Canadian Enhancements (ICD-10-CA) that were reported in DAD or NACRS within five years before lung cancer diagnosis were used to calculate a weighted Charlson comorbidity index, using the *comorbidity* package in R.[50] All diagnoses, in addition to the main diagnosis, linked to a hospitalization or visit were included in the calculation, except for lung cancer diagnoses. Individuals not found in DAD and NACRS were assigned a Charlson comorbidity score of 0.

Statistical analyses

All analyses were conducted using the statistical software R.[51] We first reported the distribution of characteristics for individuals in our study samples (i.e., frequencies, proportions, medians, and interquartile ranges). All frequencies were rounded to the nearest five, and frequencies of less than 10 were not reported as per Statistics Canada guidelines to protect data confidentiality. We then applied intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy (I-MAIHDA) to assess intersectional inequalities. Our analyses followed a recent tutorial for I-MAIHDA by Evans et al.[52]

We first obtained the frequencies of each intersectional stratum to estimate the proportion of strata with a total of individuals less than 10, which is the minimum sample size for obtaining precise and accurate results.[52] For our main analyses, two multilevel logistic regressions were fit with the package *lme4*, in which individuals at the first level are clustered within intersectional strata at the second level. First, a null model was fit, which only included random intercepts for each stratum (model 1A, equation 1), and then a main effects model was fit in which we added additive fixed effects at the second level for age, sex, income, and immigration status (model 1B, equation 2). In these models, the 1-year survival outcome y_{ij} for individual i and stratum j is assumed to

follow a Bernoulli distribution with a probability π_j (i.e., $y_{ij} \sim \text{Bernoulli}\ (\pi_j)$). The random effect from stratum j, u_j, is assumed to be normally distributed with a mean of 0 and variance σ_u^2 (i.e., $u_j \sim N(0, \sigma_u^2)$).

$$log\left(\frac{\pi_j}{1-\pi_j}\right) = \beta_0 + u_j \quad [1]$$

$$log\left(\frac{\pi_j}{1-\pi_j}\right) = \beta_0 + \sum_{z=2}^{5} \beta_{AGEe_z} \cdot AGE_{zj} + \beta_{FEM} \cdot FEM_j + \sum_{z=2}^{3} \beta_{INC_z} \cdot INC_{zj} + \beta_{IMMIG} \cdot IMMIG_j + u_j \quad [2]$$

The random intercepts in the null model allow us to capture the variance in the individual expected outcome (i.e. on the log-odds scale) that is due to strata membership (i.e., between-stratum variance). In the main effects model, the interpretation of the between-stratum variance changes to the variation in the individual expected outcome in each stratum that remains after accounting for variation that is explained by additive effects. As such, the random intercepts in the second model capture the intersectional effects that lead to outcomes that are above or below the expected additive effects of the four social identities specific to each stratum.

For all models, we obtained summary measures: the between-stratum variance in the expected outcome (i.e., σ_u^2 on the log-odds scale), the variance partition coefficient (VPC), the proportional change in variance (PCV), the median odds ratio (MOR), and the area under the receiver operating characteristic curve (AUC). Bootstrapping, in which case resampling was conducted within each stratum, was applied to obtain normal 95% confidence intervals (95% CIs) for all estimated summary measures.

The VPC in equation 3, represents the proportion of total variance in the expected outcome due to between-stratum differences. Total variance is calculated by adding the between-stratum variance to a preset individual-level variance (σ_e^2) of ~3.29, as per standard logistic distribution.[53]

$$VPC = \frac{\sigma_u^2}{\sigma_e^2 + \sigma_u^2} = \frac{\sigma_u^2}{3.29 + \sigma_u^2} \quad [3]$$

The PCV in equation [4], represents the percent change in between-stratum variance from the null model to the main effects model due to the inclusion of additive effects.

$$PCV = \frac{\sigma^2_{u\,(null\ model)} - \sigma^2_{u\,(main\ effects\ model)}}{\sigma^2_{u\,(null\ model)}} \quad [4]$$

The MOR represents the median of the distribution of odds ratios for the deviation from additive effects (intercept or intercept + main effects) across all possible social strata contrasts.[53] Each stratum's deviation from the additive effects in odds is compared to another stratum's deviation, and expressed as an odds ratio. The median odds ratio is derived from all possible paired comparisons, with the stratum having higher odds always in the numerator. The MOR quantifies the extent of differences between strata, with an MOR of 1 indicating no between-stratum differences. It is calculated using equation 5, where Φ denotes the standard normal cumulative distribution function.

$$MOR = e^{\left(\sqrt{2\,\sigma^2_u}\,\Phi(0.75)\right)} \quad [5]$$

The AUC measures each model's performance in distinguishing between individuals who survived and died within the 1-year follow-up, given their predicted 1-year survival probabilities, which are calculated with the inverse logit function. Model AUCs can range from 0.5 to 1, representing models that make predictions no better than random guessing and models that make perfect predictions. For each model, we estimated the AUC based on predictions obtained from the fixed plus random portions of the model.

We obtained stratum-specific estimates for each main effects model, including total predicted probabilities and their deviations (i.e., intersectional effects), which represent the deviation in total predicted probability above or below what is expected from predictions based solely on the fixed effects of each social identity (i.e., additive effects). These estimates (i.e., total predicted probabilities and intersectional effects) were derived by applying the inverse logit function to the model coefficients, as outlined in equations 6, 7, and 8.

$$\text{Total predicted probability} = \frac{1}{1+e^{-\left(\beta_0 + \sum_{z=2}^{5}\beta_{AGEe_z}\cdot AGE_{zj} + \beta_{FEM}\cdot FEM_j + \sum_{z=2}^{3}\beta_{INC_z}\cdot INC_{zj} + \beta_{IMMIG}\cdot IMMIG_j + u_j\right)}} \quad [6]$$

$$\text{Additive effect} = \frac{1}{1+e^{-\left(\beta_0 + \sum_{z=2}^{5}\beta_{AGEe_z}\cdot AGE_{zj} + \beta_{FEM}\cdot FEM_j + \sum_{z=2}^{3}\beta_{INC_z}\cdot INC_{zj} + \beta_{IMMIG}\cdot IMMIG_j\right)}} \quad [7]$$

$$\text{Intersectional effect} = \text{total predicted probability} - \text{additive effect} \quad [8]$$

Simulation methods were used to estimate the variance of the stratum-specific predicted probabilities. First, uncertainty in both the fixed and random effects from the main effects model was incorporated into simulations, independently, to predict stratum-specific total probabilities and their normal 95% CIs. For each simulated stratum, the probability based solely on the addition of main effects (i.e., fitted value of additive effect) was also subtracted from the total predicted probability to calculate the intersectional effect. The mean and standard deviation of the intersectional effects were used to construct 95% CIs around the mean. Caterpillar plots were used to visually present the stratum-specific total predicted probabilities and intersectional effects.

Lung cancer survival varies by province,[2] by place of residence in an urban-rural spectrum,[6] and has improved over time.[46, 47] We therefore accounted for geographical context and temporal trends in our analyses by fitting nine additional models. We first fit three cross-classified models,[54] with individuals at the first level nested within intersectional strata at the second level and either within time periods (null model 2A), province (null model 2B), or urban-rural residence (null model 2C) also at the second level. We compared the summary measures for intersectional strata in these models to those from null model 1A, to assess the importance of contextual factors. To assess variation in stratum-specific results, we also fit null and main effects models that included intersectional strata further stratified by either place of residence (models: null 2D and 2E; main effects 2F and 2F) or time period (models: null 2H; main effects 2I). In the stratified analyses, provinces were grouped into three categories based on previously reported lung cancer survival rates and observed patterns of survival in our data for territories.[2]

For our secondary analyses, we assessed the impact of additional variables for which information was available only 2010 onwards (e.g., stage at diagnosis) on the stratum-specific estimates. We first repeated the main analysis among the subset sample (models: null 3A; main effects 3B). Three separate models we also fit, which included fixed effects for comorbidities (model 3C), tumour histology (model 3D), and locally advanced or metastatic cancers (model 3E) in a stepwise fashion.

Results

Cohort characteristics

Of 193,530 individuals identified with a primary diagnosis of lung cancer in the CCR between 2002 and 2013, we included 183,350 (94.7%) in the main analyses after excluding those with missing data (Figure 6.2.1). For our secondary analyses, we retained 54,735 (74.1%) of the 71,035 individuals diagnosed between 2010 and 2013.

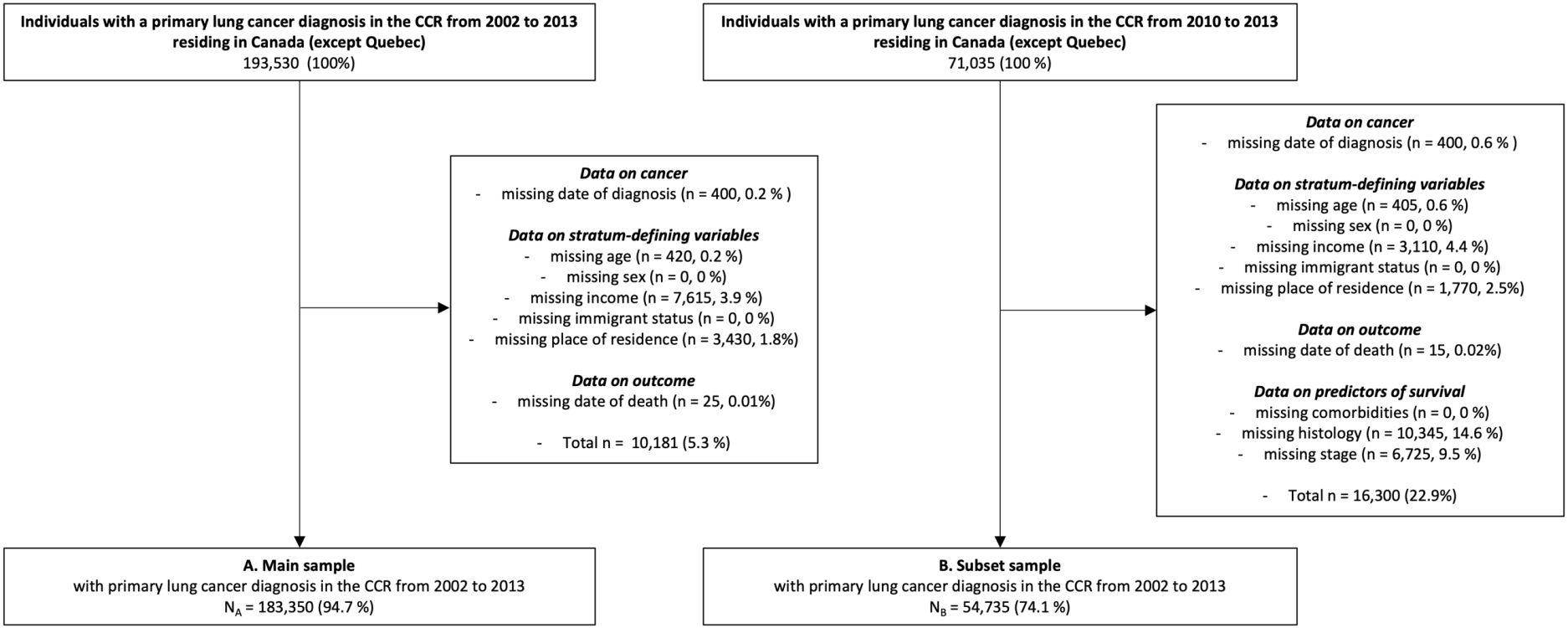


Figure 6.2.1. Flowchart of study population for main and secondary analyses

Main analyses

In our main study sample, the median age at lung cancer diagnosis was 71 years, with 47.2% of individuals being women, 50% being low-income, and 4.9% being immigrants who landed in Canada after 1980 (Table 6.2.1). The distribution of individual characteristics was similar in the subset sample, though there was a slight decrease in the percentage of individuals who were ≥80 years or in the low-income bracket. The overall 1-year survival probability increased from 39.3% in the main sample to 45.5% in the subset sample. None of the 72 intersectional strata had fewer than 10 observations in the main sample, whereas 4 strata (5.6%) did in the subset sample (Table 6.2.2 and Tables 6.3.1 and 6.3.2 in Appendix B).

Table 6.2.1. Characteristics of individuals in the CCR diagnosed with a primary lung cancer

Characteristics	**Main sample** [2002-2013, N=183,350] **n (%)**	**Subset sample** [2010-2013, N=54,735] **n (%)**
Age in years		
< 60	31,850 (17.4)	9,575 (17.5)
60-64	22,775 (12.4)	7,425 (13.6)
65-69	28,795 (15.7)	9,335 (17.1)
70-74	32,060 (17.5)	9,790 (17.9)
75-79	30,625 (16.7)	8,885 (16.2)
≥ 80	37,240 (20.3)	9,725 (17.8)
Median (interquartile range)	71 (63-78)	70 (63-77)
Female	86,455 (47.2)	26,740 (48.8)
Family income tercile [A]		
1: High	32,245 (17.6)	9,890 (18.1)
2: Medium	59,500 (32.5)	18,445 (33.7)
3: Low	91,605 (50.0)	26,400 (48.2)
Immigrant status		
Long-term resident	174,445 (95.1)	51,400 (93.9)
Immigrant (landed in 1980 or after)	8,900 (4.9)	3,335 (6.1)
Year of lung cancer diagnosis		
2002	13,380 (7.3)	
2003	13,680 (7.5)	
2004	14,115 (7.7)	
2005	14,710 (8.0)	
2006	14,750 (8.0)	
2007	15,105 (8.2)	
2008	15,220 (8.3)	
2009	15,670 (8.5)	
2010	16,220 (8.8)	13,325 (24.3)
2011	16,275 (8.9)	13,410 (24.5)
2012	17,140 (9.3)	13,935 (25.5)
2013	17,085 (9.3)	14,065 (25.7)

Table 6.2.1 (continued). Characteristics of individuals in the CCR diagnosed with a primary lung cancer

Characteristics	Main sample [2002-2013, N=183,350] n (%)	Subset sample [2010-2013, N=54,735] n (%)
Province		
Alberta	20,305 (11.1)	6,660 (12.2)
British Columbia	31,675 (17.3)	9,710 (17.7)
Manitoba	9,790 (5.3)	2,660 (4.9)
New Brunswick	7,915 (4.3)	2,355 (4.3)
Nova Scotia	9,915 (5.4)	2,840 (5.2)
Prince Edward Island	1,355 (0.7)	400 (0.7)
Ontario	89,295 (48.7)	26,000 (47.5)
Saskatchewan	8,235 (4.5)	2,550 (4.7)
Newfoundland and Labrador	4,315 (2.4)	1,400 (2.6)
Yukon, Northwest Territories or Nunavut	550 (0.3)	160 (0.3)
Urban-rural residence		
Large urban cities	107,895 (58.8)	33,200 (60.7)
Small urban cities	32,950 (18.0)	9,280 (17.0)
Rural areas and territories	42,505 (23.2)	12,255 (22.4)
Weighted Charlson comorbidity index [B]		
0		27,300 (49.9)
1-2		14,200 (25.9)
3-4		4,485 (8.2)
>=5		27,300 (16.0)
Lung cancer histology		
Adenocarcinoma	58,245 (31.8)	23,795 (43.5)
Squamous Cell Carcinoma	30,950 (16.9)	10,705 (19.6)
Large Cell Carcinoma	3,870 (2.1)	685 (1.3)
NSCLC, NOS	4,5100 (24.6)	12,505 (22.8)
Small Cell Carcinoma	20,770 (11.3)	7,030 (12.8)
Sarcomas and Other Specified	290 (0.2)	15 (0.03)
Missing	24,120 (13.2)	0 (0.0)
Lung cancer stage		
Stage I		10,550 (19.3)
Stage II		4,745 (8.7)
Stage III		11,060 (20.2)
Stage IV		28,375 (51.8)
Survived until end of 1-year follow-up	72,085 (39.3)	24,915 (45.5)

[A] After-tax family income was measured at 2 years prior to lung cancer diagnosis and equivalized to account for household size. It was classified into nationally representative terciles, specific to province, urban-rural residence (CMA/CA/Rural/Territories), and calendar year.

[B] Comorbidities measured in the 5 years prior to lung cancer diagnosis.

Abbreviations: CCR: Canadian Cancer Registry.

Table 6.2.2. Summary of stratum frequencies

Strata frequency	Main sample [2002-2013, N=183,350]		Subset sample [2010-2013, N=54,735]	
	Number of strata out of 72	**Percent of total strata**	**Number of strata out of 72**	**Percent of total strata**
less than 10	0	0.0	4	5.6
at least 10	72	100.0	68	94.4
at least 20	71	98.6	63	87.5
at least 30	69	95.8	60	83.3
at least 50	62	86.1	55	76.4
at least 100	57	79.2	52	72.2

In our null model 1A, we observed a between-strata variance of 0.20 (95% CI: 0.15, 0.24), which represents 5.8% of the total variance (VPC) in 1-year survival attributable to inequalities across intersectional strata (Table 6.2.3). The MOR in the null model indicated that there is a 50% chance that a randomly selected strata will have a deviation that is 1.53 (95% CI: 1.46, 1.61) times that of the strata with the smaller deviation. After adjusting for main effects (model 1B), the between-strata variance decreased substantially resulting in a VPC of only 0.1% and a MOR of 1.06 (95% CI: 1.01, 1.11). The PCV of 98.2% indicates that most of the differences between strata are explained by additive effects.

Table 6.2.3. Main analyses using I-MAIHDA: model estimates and summary measures for intersectionality between age, sex, income, and immigration status

	Null model 1A	ME model 1B	Cross-classified null model 2A	Cross-classified null model 2B	Cross-classified null model 2C
Predictors	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)
Intercept [baseline odds, not OR]	0.81 (0.73, 0.90)	0.89 (0.84, 0.95)	0.81 (0.72, 0.92)	0.79 (0.70, 0.89)	0.80 (0.71, 0.90)
Age in years					
< 60		1.00			
60-64		0.94 (0.87, 1.00)			
65-69		0.87 (0.81, 0.93)			
70-74		0.78 (0.73, 0.84)			
75-79		0.64 (0.60, 0.69)			
>= 80		0.42 (0.39, 0.45)			
Sex					
Male		1.00 --			
Female		1.44 (1.38, 1.50)			
Family income tercile					
1: High		1.00 --			
2: Medium		0.86 (0.82, 0.91)			
3: Low		0.72 (0.68, 0.76)			
Immigrant status					
Long-term resident		1.00 --			
Immigrant		1.59 (1.50, 1.68)			
Model summary measures					
Individual-level variance (σ^2_e)	3.29	3.29	3.29	3.29	3.29
Stratum-level variance (σ^2_u)	0.20 (0.15, 0.24)	0.0036 (-0.0035, 0.011)	0.20 strata (0.16, 0.25) 0.0053 province (0.00073, 0.0098)	0.20 strata (0.16, 0.24) 0.0019 urban-rural (0.00092, 0.0029)	0.20 strata (0.15, 0.24) 0.010 year (0.0079, 0.012)
VPC in %	5.8 (5.5, 7.0)	0.1 (0.1, 0.3)	5.8 strata (4.6, 7.0) 0.2 province (0.02, 0.3)	5.7 strata (4.5, 6.8) 0.06 urban-rural (0.03, 0.09)	5.7 strata (4.5, 6.8) 0.3 year (0.2, 0.3)
PCV in %		98.2 (95.3, 101.1)	- 0.7 (-1.3, -0.13)	1.8 (1.2, 2.4)	1.4 (0.68, 2.2)

Table 6.2.3 (continued). Main analyses: model estimates and summary measures

	Null model 1A	ME model 1B	Cross-classified null model 2A	Cross-classified null model 2B	Cross-classified null model 2C
Model summary measures					
MOR	1.53 (1.46, 1.61)	1.06 (1.01, 1.11)	1.54 strata (1.46, 1.61) 1.07 province (1.04, 1.10)	1.53 strata (1.46, 1.60) 1.04 urban-rural (1.03, 1.05)	1.53 strata (1.52, 1.60) 1.10 year (1.09, 1.11)
AUC	0.608 (0.606, 0.611)	0.608 (0.606, 0.611)	0.609 (0.606, 0.612)	0.609 (0.606, 0.611)	0.611 (0.609, 0.614)
Number of strata	72		72 strata 10 province	72 strata 3 urban-rural	72 strata 12 year

Study period 2002-2013, N =183,350.

Abbreviations: I-MAIHDA: intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy; ME: Main Effects; OR: odds ratio; 95% CI: 95% confidence interval; VPC: variance partition coefficient, PCV: percent change in variance; MOR: median odds ratio; AUC: area under the receiver operating curve.

The predicted 1-year survival probabilities from the main effect model 1B ranged from 23.3% to 66.9% across intersectional strata (Figure 6.2.2 and Table 6.3.1 in Appendix B). Although the strata with the lowest probabilities predominantly included individuals aged ≥80, and the strata with the highest probabilities mainly included individuals aged <65, some notable patterns emerged. For example, despite being <60 years old, only 36.3% of low-income long-term resident men (strata 1131) survived up to one year. This survival probability was the lowest among all other strata that included individuals of the same age and was similar to that of low-income long-term resident men aged 70–74 (33.9%) or even high-income immigrant men aged ≥80 (37.3%). Immigrant women aged <70 with medium to high incomes had the highest survival probabilities that were almost all above 60%, with low-income immigrant women aged <60 (stratum 1232) also having the 5th highest predicted survival probability at 62.6%.

The differences in predicted probabilities due to intersectional effects, as shown in Figure 2B (Table 6.3.1 in Appendix B), indicate that long-term resident men aged <60 who had the lowest or highest income (strata 1131 and 1111) and low-income immigrant women aged ≥80 (stratum 6232), had 1-year survival probabilities that were lower than expected based solely on the additive effects of age, sex, income, and immigration status. For example, the predicted probability for high-income long-term resident men aged <60 (stratum 1131) was 2.8 percentage points (95% CI: -5.0 to -0.5) lower than expected from additive effects alone. In contrast, low-income long-term resident men aged ≥80 (stratum 6131) and low-income immigrant women aged <60 (stratum 1232) had higher than expected predicted survival probabilities. Figure 6.2.3 displays the strata IDs ranked first by sex, then by immigrant status, followed by age and income, with the total predicted probabilities on the y-axis. It highlights in red or green the strata that deviated from the expected 1-year survival probabilities based solely on additive effects. Interestingly, both for low-income long-term resident men and low-income immigrant women, we observed that the intersectional effects mirrored the direction of the survival advantage or disadvantage at younger ages but reversed at older ages. For example, long-term resident men showed the lowest survival probabilities compared to all other strata within each age group, including those aged 80 and older. However, this probability was higher than expected based solely on additive effects (stratum 6131 in green).

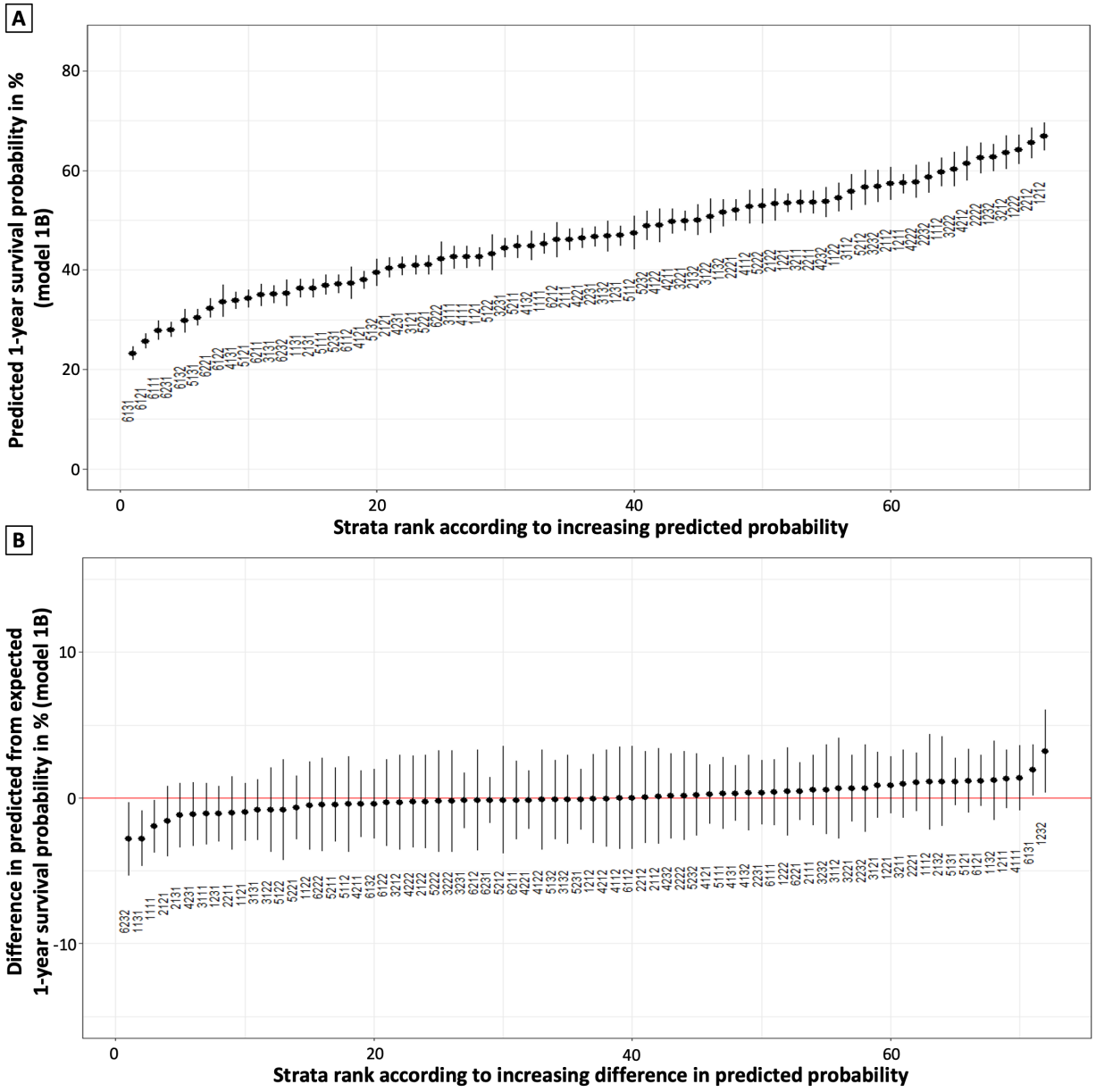


Figure 6.2.2. Main analysis: total predicted probabilities of 1-year lung cancer survival (A) and differences in total predicted probabilities from the expected probabilities based solely on fixed effects (B).

Study period 2002-2013, N =183,350. Estimates were obtained from main effects model 1B. The estimate in panel B represents the stratum-specific intersectional effect.

Strata ID defined with 4-digit codes with:

- the 1st position representing age in years (1: <60; 2: 60-64; 3: 65-69; 4: 70-74; 5: 75-79; and 6: ≥80)
- the 2nd position representing sex (1: male; and 2: female)
- the 3rd position representing family income tercile (1: high income, 2: medium income, and 2: low income)
- the 4th position representing immigration status (1: long-term resident; and 2: immigrant)

For example, the stratum "1231" represents women aged <60 who are low-income and long-term residents of Canada.

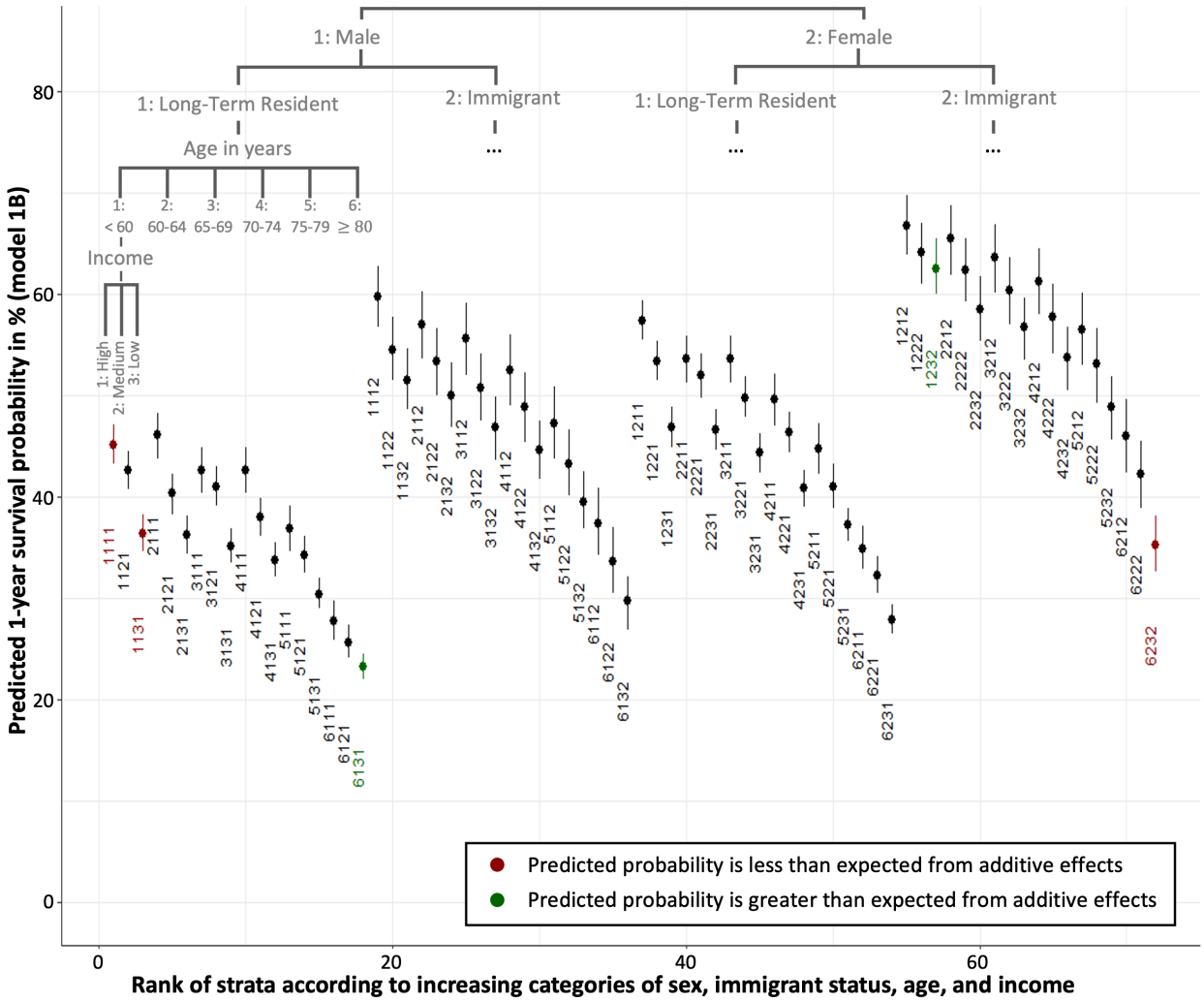


Figure 6.2.3. Predicted probabilities of 1-year lung cancer survival, highlighting strata deviating from expected probabilities based solely on fixed effects.

Study period 2002-2013, N =183,350. Estimates were obtained from main effects model 1B.

The strata IDs are ranked according to increasing order of the category labels of sex, immigrant status, age, and income. This means strata 1111 is ranked first and 6232 is ranked last.

Strata ID defined with 4-digit codes with:

- the 1st position representing age in years (1: <60; 2: 60-64; 3: 65-69; 4: 70-74; 5: 75-79; and 6: ≥80)
- the 2nd position representing sex (1: male; and 2: female)
- the 3rd position representing family income tercile (1: high income, 2: medium income, and 2: low income)
- the 4th position representing immigration status (1: long-term resident; and 2: immigrant)

For example, the stratum "1231" represents women aged <60 who are low-income and long-term residents of Canada.

Cross-classified and stratified analyses

The cross-classified models 2A, 2B, and 2C, which incorporated additional contextual factors into the analysis, showed that clustering at the levels of province, urban-rural residence, and year of diagnosis accounted for only 0.2%, 0.06%, and 0.3% of the total variance in 1-year survival, respectively (VPCs in Table 6.2.3). Furthermore, the VPC for intersectional strata attenuated only slightly from 5.8% in the null model 1A to 5.7% in the cross-classified models that included urban-rural residence or year of diagnosis as second-level strata (models 2B and 2C). Overall, this suggests that clustering by time, province, and urban-rural area- in that order- is important, but it is negligible compared to clustering by intersectional strata.

Stratifying the intersectional strata by province (3 categories), urban-rural residence, and time period (3 categories) resulted in VPCs of 5.7%, 5.4%, and 6.0% in null models 2D, 2F, and 2H, respectively (Table 6.2.4). It is important to note that the percentage of strata with fewer than 10 observations was 16.8% and 18.5% in models 2D and 2F, respectively, which may have led to shrinkage of the random intercepts and attenuation of the VPC. Despite 5% of strata having frequencies of less than 10 due to stratification by time period, we observed a slightly higher VPC of 6.0% in null model 2H compared to the null model 1A. Overall, including main effects in the stratified models (models 2E, 2G, and 2I) resulted in PCVs of approximately 98%. This means that about 2% of the observed variation in 1-year survival observed between strata could not be explained by additive effects alone in these models, an equivalent result to that found in the main analysis (model 1B).

Table 6.2.4. Stratified analyses using I-MAIHDA: model estimates and summary measures for intersectionality between age, sex, income, and immigration status

	Stratified by province		Stratified by urban-rural residence		Stratified by diagnosis year	
	Null model 2D	**ME model 2E**	**Null model 2F**	**ME model 2G**	**Null model 2H**	**ME model 2I**
Predictors	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)
Intercept [baseline odds, not OR]	0.76 (0.72, 0.81)	0.82 (0.78, 0.86)	0.75 (0.70, 0.80)	0.91 (0.86, 0.95)	0.80 (0.75, 0.86)	0.79 (0.75, 0.83)
Age in years						
< 60		1.00 --		1.00 --		
60-64		0.95 (0.90, 0.99)		0.95 (0.90, 1.00)		1.00 --
65-69		0.88 (0.84, 0.93)		0.89 (0.84, 0.93)		0.93 (0.88, 0.98)
70-74		0.80 (0.76, 0.84)		0.80 (0.76, 0.84)		0.88 (0.83, 0.92)
75-79		0.67 (0.64, 0.70)		0.66 (0.63, 0.70)		0.79 (0.75, 0.83)
>= 80		0.44 (0.42, 0.46)		0.44 (0.42, 0.46)		0.66 (0.62, 0.69)
Sex						
Male		1.00 --		1.00 --		1.00 --
Female		1.44 (1.40, 1.48)		1.42 (1.38, 1.47)		1.43 (1.39, 1.47)
Family income tercile						
1: High		1.00 --		1.00 --		1.00 --
2: Medium		0.88 (0.84, 0.91)		0.87 (0.84, 0.91)		0.87 (0.83, 0.90)
3: Low		0.72 (0.69, 0.75)		0.73 (0.70, 0.75)		0.72 (0.69, 0.74)
Immigrant status						
Long-term resident		1.00 --		1.00 --		1.00 --
Immigrant		1.61 (1.53, 1.69)		1.55 (1.47, 1.63)		1.57 (1.49, 1.65)
Province [A]						
BC		1.00 --				
MN and AB		1.08 (1.03, 1.13)				
ON, NB, NFL		1.09 (1.05, 1.13)				
SK, NS, PEI, and YK/NWT/NU		1.01 (0.96, 1.06)				
Urban-rural residence						
Large urban cities				1.00 --		
Small urban cities				0.91 (0.87 – 0.94)		
Rural areas				0.93 (0.90 – 0.96)		
Year of diagnosis						
2002-2004						1.00 --
2005-2010						1.16 (1.12, 1.20)
2011-2013						1.31 (1.26, 1.36)

Table 6.2.4 (continued). Stratified analyses: model estimates and summary measures

	Stratified by province		Stratified by urban-rural residence		Stratified by diagnosis year	
	Null model 2D	**ME model 2E**	**Null model 2F**	**ME model 2G**	**Null model 2H**	**ME model 2I**
Model summary measures						
Individual-level variance (σ_e^2)	3.29	3.29	3.29	3.29	3.29	3.29
Stratum-level variance (σ_u^2)	0.20 (0.13, 0.27)	0.0043 (-0.0096, 0.018)	0.19 (0.11, 0.26)	0.0031 (-0.0059, 0.012)	0.21 (0.14, 0.28)	0.0037 (-0.0078, 0.015)
VPC in %	5.7 (3.7, 7.6)	0.1 (-0.3, 0.6)	5.4 (3.4, 7.4)	0.1 (-0.2, 0.4)	6.0 (4.1, 7.8)	0.1 (-0.002, 0.005)
PCV in %	---	97.9 (92.5, 103.2)	---	98.3 (94.8, 101.9)	---	98.2 (96.8, 98.3)
MOR	1.53 (1.41, 1.64)	1.06 (1.09, 1.12)	1.51 (1.39, 1.63)	1.05 (0.99, 1.11)	1.55 (1.44, 1.66)	1.06 (0.99, 1.13)
AUC	0.610 (0.606, 0.615)	0.609 (0.605, 0.613)	0.610 (0.607, 0.612)	0.609 (0.606, 0.611)	0.612 (0.610, 0.615)	0.611 (0.609, 0.614)
Number of strata	285 [B]		216		216	
Number of strata with frequency ≥ 20 (%)	215 (75.4%)		147 (68%)		184 (85.2%)	
Number of strata with frequency < 10 (%)	48 (16.8%)		40 (18.5%)		11 (5.0%)	

Study period 2002-2013, N =183,350.

[A] Provinces grouped according to previously reported 1-year survival rates.[2]

[B] The following strata had a frequency of 0: 4112 in SK, NS, PEI, and YK/NWT/NU, 5212 in SK, NS, PEI, and YK/NWT/NU, 6212 in MN and AB.

Abbreviations: ME: Main Effects; OR: odds ratio; 95% CI: 95% confidence interval; VPC: variance partition coefficient, PCV: percent change in variance; MOR: median odds ratio; AUC: area under the receiver operating curve, BC: British Columbia; MN: Manitoba; AB: Alberta; ON: Ontario; NB: New Brunswick; NFL: Newfoundland and Labrador; SK: Saskatchewan; NS: Nova Scotia; PEI: Prince Edward Island; YK: Yukon; NWT: Northwest Territories; NU: Nunavut.

For the analyses stratified by place of residence (models 2E and 2G) and time period (model 2I), we also examined stratum-specific estimates (Appendix C). We focused on the strata that exhibited significant intersectional effects in the non-stratified analysis (model 1B), which included low-income long-term resident men aged <60 or ≥80 (strata 1131 and 6131), high-income long-term resident men aged <60 (stratum 1111), and low-income immigrant women aged <60 or ≥80 (strata 1233 and 6232). Figure 6.2.4 illustrates that the intersectional effects for high-income long-term resident men aged <60 (stratum 1111) were most pronounced in rural areas and during the period 2005-2010. In contrast, the intersectional effects for low-income long-term resident men aged <60 (stratum 1131) were most pronounced in Ontario, New Brunswick, or Newfoundland. Both the youngest and oldest strata of low-income long-term resident men (strata 1131 and 6131) showed the most pronounced intersectional effects in the most recent period (2011-2013). For the youngest and oldest low-income immigrant women (strata 1232 and 6232), the intersectional effects were nearly absent in the most recent period.

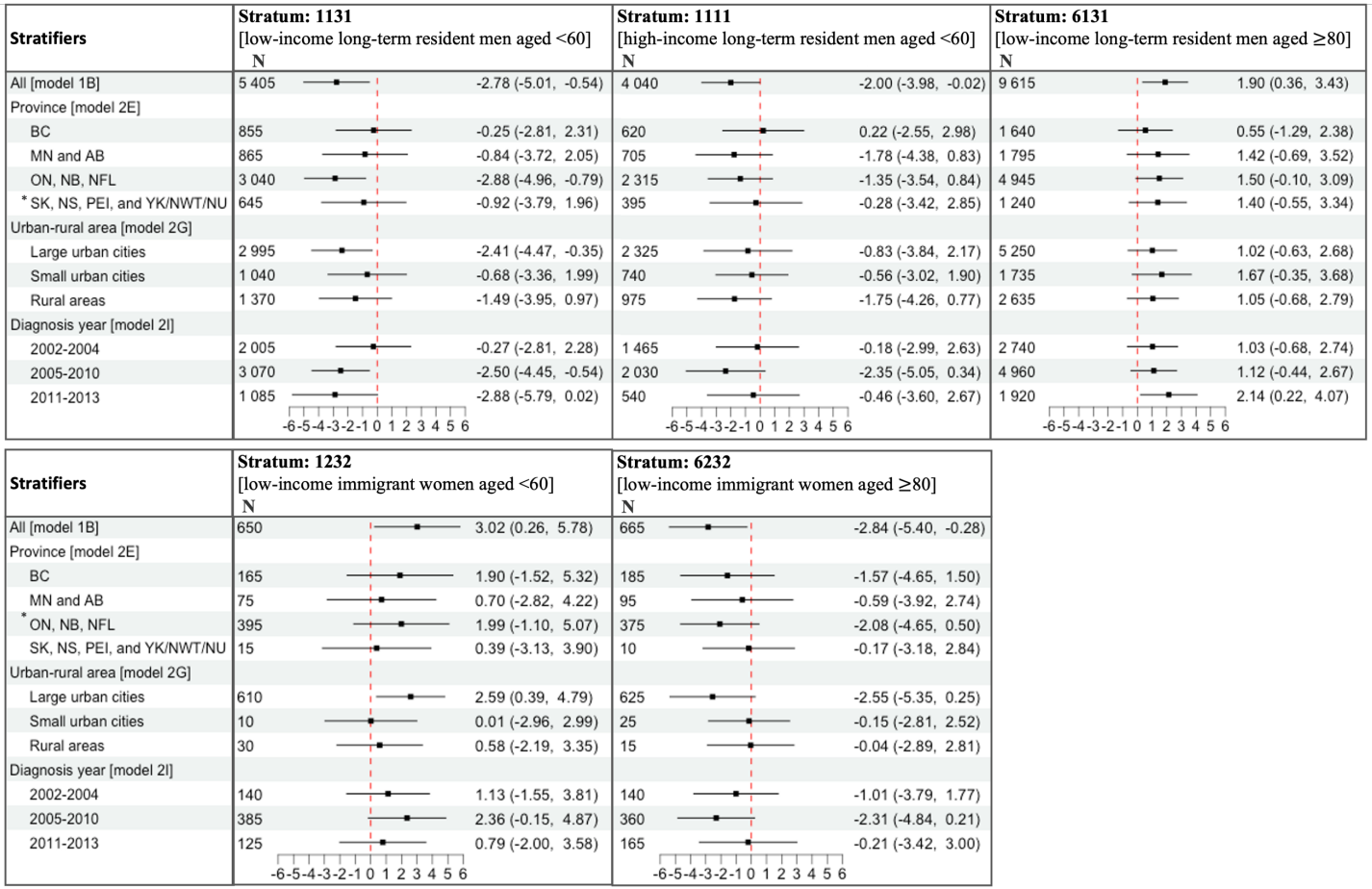


Figure 6.2.4. Stratified analyses: stratum-specific differences in predicted 1-year survival probabilities from the expected probabilities based solely on fixed effects.

Study period 2002-2013, N =183,350. Estimates were obtained from main effects models 1B, 2E (stratification by province), 2G (stratification by urban-rural area), and 2I (stratification by time period).

* Provinces grouped according to previously reported 1-year survival rates.[2]

Abbreviations: BC: British Columbia; MN: Manitoba; AB: Alberta; ON: Ontario, NB: New Brunswick, NFL: Newfoundland and Labrador; SK: Saskatchewan; NS: Nova Scotia; PEI: Prince Edward Island; YK: Yukon; NWT: Northwest Territories; NU: Nunavut.

Secondary analyses

In our secondary analyses, which included only individuals diagnosed from 2010 to 2013, the VPC in the null model decreased to 4.7 % (model 3A in Table 6.2.5) from 5.7% in the main analysis (model 3A in Table 6.2.5). As in the main analysis, the addition of main effects in model 3B reduced the VPC from model 3A by 98.7% and changed the MOR from 1.47 (95% CI: 1.36, 1.58) to 1.04 (95% CI: 0.97, 1.12). When comorbidities and histology were further included, the AUC improved from 0.591 in model 3B to 0.643 and 0.677 in models 3C and 3D, respectively, while the point estimates for the MOR increased slightly from 1.04 to 1.05. This suggests that comorbidities and histology accounted more for individual-level variability, which likely reduced shrinkage of the random effects and increased variability between strata. When the stage at diagnosis was included in model 3E, the AUC increased to 0.836, and the point estimate MOR decreased to 1.02, with nearly all of the between-strata variance explained by additive main effects (PCV = 99.2%).

The intersectional effects in Figure 6.2.5 (Table 6.3.3 in Appendix D), observed from model 3B were attenuated compared to those from the main analysis (model 1B in Figure 6.2.2). The largest deviations in predicted probabilities remained for the youngest and oldest low-income long-term resident men (strata 1131 and 6133) but were no longer statistically significant. In contrast, intersectional effects for high-income long-term resident men aged <60 (stratum 1111) and low-income immigrant women aged <60 or ≥80 (strata 1232 and 6232) were attenuated to less than 1%, which is congruent with our stratified analyses by time period (Figure 6.2.4). Figure 6.2.5 and Table 6.2.6 also show that further adjustments with comorbidities and histology (models 3C and 3D) did not have a large effect on the intersectional effects. However, adjusting for stage at diagnosis (model 3E) reduced intersectional effects to less than 1%.

Table 6.2.5. Secondary analyses using I-MAIHDA: model estimates and summary measures for intersectionality between age, sex, income, and immigration status

	Null model 3A	ME model 3B	Comorbidity-adjusted ME model 3C	Histology-adjusted ME model 3D	Stage-adjusted ME model 3E
Predictors	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)
Intercept [baseline odds]	1.00 0.90 – 1.11	0.98 0.92 – 1.06	1.18 1.09 – 1.27	1.56 1.45 – 1.68	19.69 17.82 – 21.75
Age in years					
< 60		1.00 --	1.00 --	1.00 --	1.00 --
60-64		0.95 0.88 – 1.03	0.95 0.87 – 1.03	0.94 0.87 – 1.02	0.86 0.79 – 0.93
65-69		0.93 0.86 – 1.01	0.93 0.86 – 1.01	0.91 0.85 – 0.99	0.79 0.73 – 0.86
70-74		0.89 0.83 – 0.96	0.89 0.82 – 0.97	0.88 0.81 – 0.95	0.69 0.64 – 0.74
75-79		0.77 0.71 – 0.83	0.76 0.70 – 0.83	0.75 0.69 – 0.81	0.58 0.53 – 0.62
>= 80		0.56 0.51 – 0.60	0.55 0.51 – 0.60	0.55 0.51 – 0.60	0.41 0.38 – 0.44
Sex					
Male		1.00 --	1.00 --	1.00 --	1.00 --
Female		1.46 1.40 – 1.53	1.45 1.38 – 1.53	1.45 1.38 – 1.52	1.41 1.35 – 1.48
Family income tercile					
1: High		1.00 --	1.00 --	1.00 --	1.00 --
2: Medium		0.87 0.82 – 0.93	0.88 0.82 – 0.94	0.89 0.84 – 0.95	0.85 0.79 – 0.90
3: Low		0.73 0.69 – 0.77	0.75 0.70 – 0.80	0.78 0.73 – 0.83	0.70 0.66 – 0.75
Immigrant status					
Long-term resident		1.00 --	1.00 --	1.00 --	1.00 --
Immigrant		1.60 1.48 – 1.73	1.54 1.42 – 1.67	1.43 1.32 – 1.55	1.81 1.66 – 1.98
Weighted Charlson Comorbidity Index					
0			1.00 --		
1-2			1.00 0.95 – 1.04	1.01 0.97 – 1.06	0.84 0.80 – 0.89
3-4			0.89 0.83 – 0.95	0.91 0.86 – 0.98	0.66 0.61 – 0.71
>=5			0.28 0.26 – 0.29	0.29 0.27 – 0.30	0.41 0.38 – 0.44
Lung cancer histology					
Adenocarcinoma				1.00 --	1.00 --
Squamous Cell Carcinoma				0.92 0.88 – 0.96	0.64 0.60 – 0.68
Large Cell Carcinoma				0.54 0.46 – 0.63	0.49 0.40 – 0.59
NSCLC, NOS				0.50 0.47 – 0.52	0.50 0.47 – 0.53
Small Cell Carcinoma				0.39 0.37 – 0.42	0.63 0.59 – 0.67
Sarcomas and Other Specified				0.75 0.27 – 2.05	0.38 0.13 – 1.10

Table 6.2.5 (continued). Secondary analyses: model estimates and summary measures

	Null model 3A	ME model 3B	Comorbidity-adjusted ME model 3C	Histology-adjusted ME model 3D	Stage-adjusted ME model 3E
Predictors	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)	1-year survival OR (95% CI)
Lung cancer stage					
Stage I					1.00 --
Stage II					0.42 0.38 – 0.46
Stage III					0.15 0.14 – 0.16
Stage IV					0.03 0.03 – 0.04
Model summary measures					
Individual-level variance (σ_e^2)	3.29	3.29	3.29	3.29	3.29
Stratum-level variance (σ_u^2)	0.16 (0.098, 0.23)	0.0021 (-0.0085, 0.013)	0.0029 (-0.0081, 0.014)	0.0022 (-0.0080, 0.012)	0.00093 (-0.011, 0.013)
VPC in %	4.7 (2.9, 6.5)	0.1 (-0.2, 0.3)	0.1 (-0.2, 0.4)	0.1 (-0.2, 0.3)	0.03 (-0.3, 0.3)
PCV in %		98.7 (93.7, 103.8)	98.2 (93.0, 103.4)	98.7 (93.8, 103.6)	99.4 (93.2, 105.6)
MOR	1.47 (1.36, 1.58)	1.04 (0.97, 1.12)	1.05 (0.98, 1.13)	1.05 (0.97, 1.12)	1.03 (0.93, 1.13)
AUC	0.592 (0.587, 0.598)	0.591 (0.586, 0.595)	0.643 (0.638, 0.647)	0.677 (0.672, 0.681)	0.836 (0.833, 0.839)
Number of strata	72				

Study period 2010-2013, N =54,735.

Abbreviations: ME: Main Effects; OR: odds ratio; 95% CI: 95% confidence interval; VPC: variance partition coefficient, PCV: percent change in variance; MOR: median odds ratio; AUC: area under the receiver operating curve.

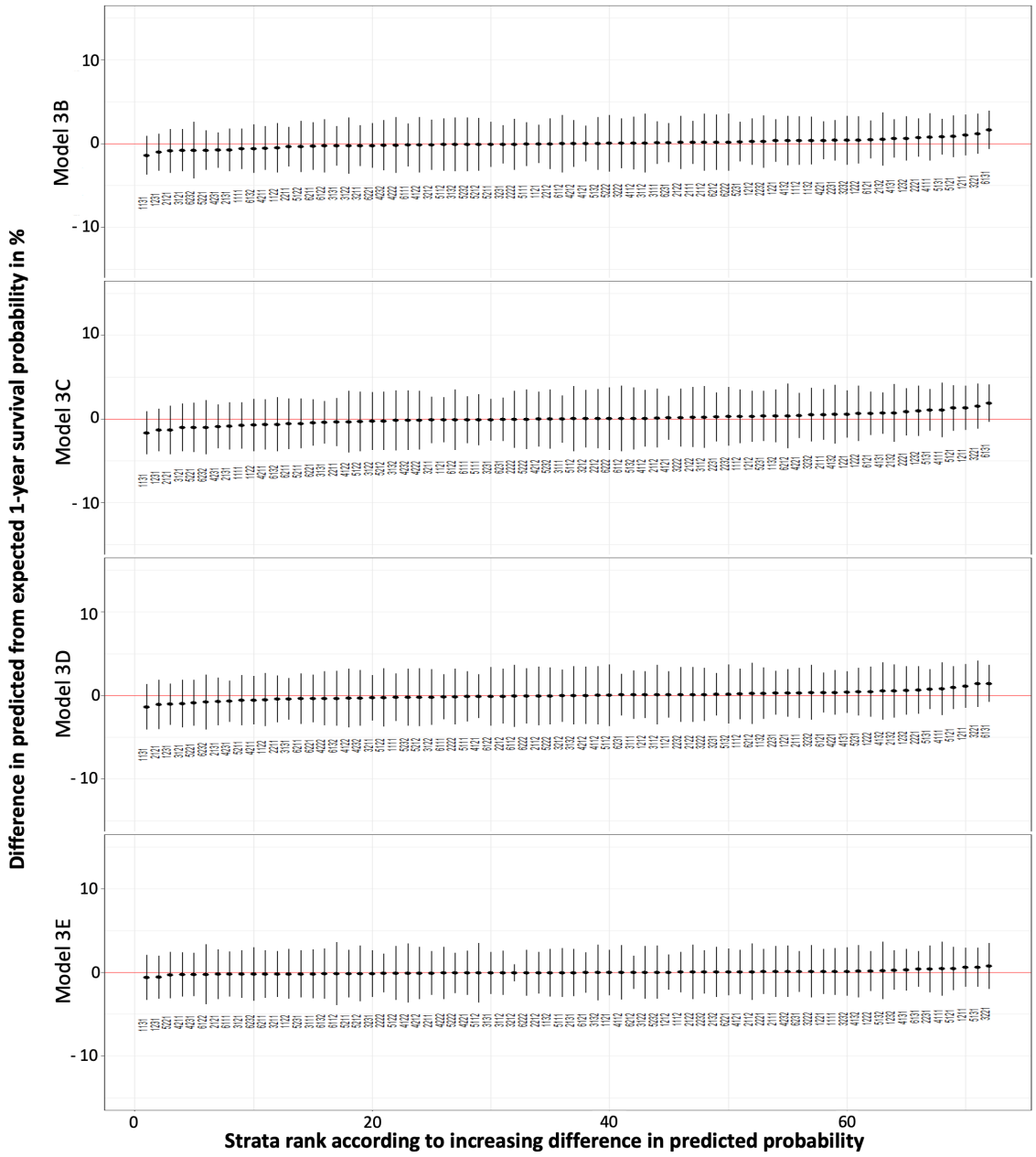


Figure 6.2.5. Secondary analyses: stratum-specific differences in predicted 1-year survival probability from the expected probabilities based solely on fixed effects.

Study period 2010-2013, N = 54,735. Estimates obtained from main effects models 3B, 3C, 3D, and 3E. Model 3B included main effects for age, sex, income, and immigrant status. An additional adjustment for comorbidities was included in model 3C, for comorbidities and histology in model 3D, and for comorbidities, histology and stage at diagnosis in model 3E.

Strata ID defined with 4-digit codes with:

- the 1st position representing age in years (1: <60; 2: 60-64; 3: 65-69; 4: 70-74; 5: 75-79; and 6: ≥80)
- the 2nd position representing sex (1: male; and 2: female)
- the 3rd position representing family income tercile (1: high income, 2: medium income, and 2: low income)
- the 4th position representing immigration status (1: long-term resident; and 2: immigrant)

For example, the stratum "1231" represents women aged <60 who are low-income and long-term residents of Canada.

Table 6.2.6. Secondary analyses: total predicted probabilities and differences from predicted probabilities based on fixed effects alone for specific intersectional strata

Stratum ID	Additional adjustments in main effects (ME) model	Model	N	Predicted 1-year survival probability in % (95% CI)	Difference in predicted from expected 1-year survival probability in % (95% CI)
6232: ≥80 years, female, low-income, immigrant	ME	3B	210	47.46 (44.37, 50.37)	-0.98 (-3.96, 2.01)
	ME + comorbidities	3C		46.10 (42.82, 49.47)	-1.08 (-4.65, 2.48)
	ME + comorbidities + histology	3D		46.56 (43.42, 49.75)	-0.85 (-3.99, 2.29)
	ME + comorbidities + histology + stage	3E		40.85 (38.67, 43.03)	-0.16 (-3.26, 2.94)
1131: <60 years, male, low-income, long-term resident	ME	3B	1 505	40.43 (38.22, 42.87)	-1.34 (-3.92, 1.23)
	ME + comorbidities	3C		40.30 (37.63, 42.67)	-1.66 (-4.50, 1.17)
	ME + comorbidities + histology	3D		40.47 (38.09, 42.85)	-1.30 (-4.04, 1.45)
	ME + comorbidities + histology + stage	3E		40.52 (38.77, 42.30)	-0.65 (-3.30, 2.01)
1111: <60 years, male, high-income, long-term resident	ME	3B	1 045	49.17 (46.66, 51.68)	-0.53 (-3.17, 2.10)
	ME + comorbidities	3C		49.15 (46.59, 51.83)	-0.71 (-3.66, 2.23)
	ME + comorbidities + histology	3D		48.61 (45.98, 51.10)	-0.25 (-3.16, 2.67)
	ME + comorbidities + histology + stage	3E		47.73 (45.66, 49.66)	0.17 (-2.70, 3.04)
6131: ≥80 years, male, low-income, long-term resident	ME	3B	2 365	30.05 (28.15, 31.90)	1.59 (-0.41, 3.59)
	ME + comorbidities	3C		30.29 (28.18, 32.50)	1.99 (-0.41, 4.38)
	ME + comorbidities + histology	3D		30.20 (28.09, 32.41)	1.41 (-0.85, 3.68)
	ME + comorbidities + histology + stage	3E		30.62 (29.05, 32.20)	0.45 (-1.69, 2.59)
1232: <60 years, female, low-income, immigrant	ME	3B	215	63.41 (60.48, 65.90)	0.58 (-2.11, 3.26)
	ME + comorbidities	3C		63.27 (60.21, 66.38)	0.90 (-2.24, 4.04)
	ME + comorbidities + histology	3D		64.53 (61.60, 67.34)	0.68 (-2.25, 3.61)
	ME + comorbidities + histology + stage	3E		64.19 (62.03, 66.36)	0.26 (-2.20, 2.72)

Study period 2010-2013, N =54,735.
Model 3B included main effects (ME) for age, sex, income, and immigrant status. An additional adjustment for comorbidities was included in model 3C, for comorbidities and histology in model 3D, and for comorbidities, histology and stage at diagnosis in model 3E.

Discussion

Among 183,350 individuals diagnosed with lung cancer in Canada between 2002 and 2013, multiple axes of advantage or disadvantage—across age, sex, income, and immigration status—primarily acted additively on 1-year survival, leading to significant variations across intersectional strata. However, 2% of the variation across intersectional strata could not be explained by additive effects alone. A closer inspection of stratum-specific deviations from predicted probabilities based on additive effects alone revealed evidence of interaction in a few strata. Significant deviations were observed among the youngest and oldest low-income long-term resident men, the youngest high-income long-term resident men, and the youngest and oldest low-income immigrant women. These intersectional effects were context-specific in terms of place of residence and time period and were most likely driven by differences in stage at diagnosis.

The lower-than-expected predicted survival probability for low-income long-term resident men aged <65 (stratum 1131) may be partly explained by smoking, as these social identities are associated with higher smoking rates.[55] Moreover, individuals who smoke at the time of their cancer diagnosis, including those with lung cancer, are more likely to exhibit these social characteristics, be diagnosed at an advanced stage, and experience lower 1-year survival.[56] Further inspection of our subset sample revealed that low-income men aged <60 had the highest rates of stage III and IV cancers (~80%) among all strata. Interestingly, despite these patterns, this group also had one of the highest rates of adenocarcinoma (~30%) among long-term resident men, a cancer type less strongly linked to smoking compared to other histologies and known for longer lung cancer survival rates.[57, 58] Men with the same characteristics but with high income (stratum 1111) also showed a similarly lower-than-expected predicted survival in the main or some of the stratified analyses. However, this intersectional effect was less prominent in more recent years of diagnosis (2011-2013), unlike low-income long-term resident men aged <65 who continued to demonstrate lower-than-expected survival probabilities (stratum 1131).

Immigrant women generally displayed one of the highest 1-year survival probabilities, including those with low income, relative to all other strata within each age group. Furthermore, those with low-income aged <60 (stratum 1232) demonstrated higher-than-expected survival probabilities. This phenomenon might stem from the impact of low income being less pronounced among immigrant women than men, largely due to traditional gender roles. The responsibility of providing

for the family tends to weigh more heavily on immigrant men than on women.[59] Conversely, as low-income immigrant women reached the age of 80 years or older (stratum 6232), they exhibited lower-than-expected survival probabilities based solely on additive effects, despite still maintaining one of the highest 1-year survival probabilities compared to other strata within the same age group. The reversing of the intersectional effect at older ages in these women may be due to a gradual diminishing of the healthy immigrant effect over time, which is more pronounced in women than in men,[60] and which may be even more pronounced in low-income women. Older low-income immigrant women, whose healthcare needs increase with rising comorbidities, may encounter greater obstacles when navigating the healthcare system relative to younger, non-immigrant women, or those with higher incomes.[61, 62]

It was challenging to draw definitive conclusions about patterns of intersectional effects by place of residence for low-income immigrant women at younger and older ages due to small cell counts and potential shrinkage of random effects. Regarding time periods, we observed no strong intersectional effects in the most recent period, possibly confounded by age-period-cohort effects. For example, our study identified immigrants arriving after 1980, meaning the most recent period may include women with the longest time since immigration, more likely to be White, and those who immigrated as middle-aged adults. Research indicates immigrant health varies by sex, age, time since immigration, and visible minority status.[63]

Overall, the number of observed intersectional effects was limited, but those identified likely reflect genuine intersectional effects, as shrinkage of random effects toward the overall mean helps guard against spurious findings due to multiple testing.[52] Furthermore, a previous Canadian study using I-MAIHDA to examine self-rated health across intersectional strata—defined by sex, income, immigrant status, and visible minority status—reported that low-income White Canadian men experience the most negative intersectional effects, whereas low-income immigrant visible-minority women exhibit some of the most positive intersectional effects.[64]

In our secondary analyses, intersectional effects were attenuated but remained qualitatively the same as those observed in the main analyses stratified by time. The attenuation could be due to shrinkage or random effects, as 5.6% of the 72 strata in the subset sample had fewer than 10 observations, or simply due to the lack of intersectional effects over time for certain strata.

Adjustment for stage at diagnosis accounted for nearly all of the excess between-strata variation. Although it may not play a large role in mediating income-based inequalities,[6, 65] our findings suggest that survival differences related to income—when considered at the intersection of age, sex, and immigration status, and beyond what would be expected from additive effects alone—may be primarily driven by stage at diagnosis. This underscores the limitations of examining inequalities along single axes of identity, rather than through an intersectional lens that recognizes the "interlocking, inseparable, and mutually constituted" nature of social advantage and disadvantage.[52]

There are limitations in our study regarding the social identities used to define strata. For immigration status, we could only identify immigrants who arrived in Canada after 1980, with a maximum time since immigration of 33 years. The median age at lung cancer diagnosis is 71 years, while the average age at immigration is 30 years.[66] To capture immigrants who migrated at age 30 and were diagnosed at 71 between 2002 and 2013, landing data from 1961–1972 would have been necessary but was unavailable. Although immigrants represent 22% of the general Canadian population,[67] they accounted for only 5% of our study cohort. Older immigrants often migrate for family reunification and may have greater support than those who migrate at younger ages, who typically arrive for economic reasons. In our previous study on lung cancer survival, where 20% of the population were immigrants—75% of whom had been in Canada for over 33 years—we found no protective effect of immigrant status. However, stratifying by sex revealed a potentially harmful effect among women. Future research should include diagnoses post-2013 to capture immigrants with longer residence and distinguish individuals by time since immigration and age at immigration to better assess intersectional effects with age, sex, and income. Incorporating race or period of immigration is also critical, given the significant increase in non-European immigration between 2001 and 2021, which has reshaped Canada's racial demographics.[68]

Another limitation is that there were very few people who lived in territories to be analyzed separately in our analyses, and Indigenous populations, who are known to have the poorest lung cancer survival rates in Canada, were not identifiable.[2, 3, 69] Additionally, we lacked information on gender, a non-binary and fluid social construct tied to structural discrimination. The observed differences by sex, a binary biological variable, may only partly reflect gender-

related disparities. Future research should incorporate gender and other intersecting social identities, such as sex, race, income, and immigration status, to better capture the complex systems influencing cancer outcomes.

Overall, despite observing intersectional effects, the null model 1A had a relatively low VPC of 5%, indicating that there remains considerable heterogeneity in 1-year survival within strata. Although observing VPCs of less than 5% is the norm for I-MAIHDA applications,[70] they indicate that discriminatory accuracy is on the lower end of the spectrum. This highlights the importance of integrating additional information when defining intersectional strata in future research, such as ethnicity, which may help uncover additional key populations facing lung cancer inequalities. Our analyses are also descriptive in nature. While we explored the potential role of certain variables, such as stage at diagnosis, in explaining intersectional effects, more formal qualitative and quantitative work (e.g., mediation analysis) is needed to understand why certain strata exhibit lower-than-expected or higher-than-expected predicted probabilities.[71] We acknowledge that adjustments for mediators can induce collider bias in the estimation of main additive effects; however, our analysis did not focus on estimating independent associations of each social identity with lung cancer survival. Instead, we aimed to understand the potential of each stratum in predicting the outcome after accounting for important explanatory variables.

Finally, we adopted a frequentist approach to obtain approximate prediction intervals through simulations, assuming a multivariate normal distribution for both fixed and random effects, while treating them as independent and not accounting for their covariance. We chose this approach because it is less computationally intensive.[52] However, this efficiency was offset by the additional cluster bootstrapping step we added to obtain confidence intervals for our model summary measures. Future work should consider the Bayesian framework, as it offers greater flexibility in modelling relationships between random and fixed effects and provides more intuitively interpretable uncertainty measures through credible intervals.

Our study also has some significant strengths. We are one of the first to quantitatively explore intersectionality based on multiple social identities in relation to cancer outcomes in Canada. Unlike traditional regression-based analyses, I-MAIHDA provides stable estimates for stratum with low frequencies by shrinking strata estimates towards a global average of the outcome. It is a

particularly more robust method for analyses with an increasing number of strata.[72] Hence, the application of I-MAIHDA in our study allowed us to assess each stratum individually, including the strata composed of older high-income immigrants that had some of the lowest frequencies. Using the CCR to create our study population, we were able to provide national-level estimates of intersectional inequalities, while taking the context of place of residence and calendar time into consideration. Furthermore, due to linkages to other databases, family income was included as an individual-level rather than an area-level measure. Finally, despite our multiple analyses, the same strata as in our main analyses often emerged as having higher- or lower-than-expected probabilities. This contributes to the robustness of our findings, especially for low-income long-term resident men.

Conclusion

In conclusion, our findings demonstrate that age, sex, income, and immigration contribute additively to 1-year lung cancer survival in Canada, resulting in significant inequalities between some intersectional strata. The application of I-MAIHDA identified potential key populations experiencing intersecting effects of these social identities. While stage at diagnosis appeared to explain much of the variation between strata attributed to intersectional effects, more data with longer follow-up and formal mediation analysis is needed to confirm its role. Future qualitative and quantitative research incorporating additional measures of structural barriers—such as healthcare access, drug coverage, treatment challenges linked to smoking stigma or financial burden, social support, and living arrangements—could provide deeper insights into how these key populations experience intersectional effects.

Funding

This research was funded by a doctoral research award (Frederick Banting and Charles Best Canada Graduate Scholarships) granted to S.Q. by the Canadian Institutes of Health Research (CIHR funding number FBD-175860). It was also supported by the Emerging Scholars Grant awarded by the Canadian Research Data Centre Network (CRDCN). E.S. holds a McGill University William Dawson Scholar award, which supports her research.

Conflict of interest

The authors declare no other conflicts of interest.

Ethics approval

McGill University's Faculty of Medicine and Health Sciences Institutional Review Board granted ethics approval this study on April 12, 2022 (study number A04-M30-22A).

Author contributions

Samia Qureshi devised the study, curated the data, conducted all analyses, drafted the manuscript and interpreted the results. **Arijit Nandi**, and **Erin Strumpf** contributed to the interpretation of the results and reviewed the manuscript. Erin Strumpf supervised the study.

Data availability

This study used Statistics Canada data accessed through a Research Data Centre (RDC) within the CRDCN. The data were part of the project that linked the Canadian Cancer Registry to multiple other administrative databases: "*Linkage of Socioeconomic and Treatment Data to the Canadian Cancer Registry*". Due to the confidential nature of these microdata, they cannot be shared directly; however, access can be requested through the CRDCN here.

Acknowledgements

The analysis presented in this paper was conducted at the Quebec Interuniversity Centre for Social Statistics (QICSS), which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the QICSS are made possible by the financial or in-kind support of the Social Sciences and Humanities Research Council (SSHRC), the Canadian Institutes of Health Research (CIHR), the Canada Foundation for Innovation (CFI), Statistics Canada, the Fonds de recherche du Québec - Société et culture (FRQSC), the Fonds de recherche du Québec -

Santé (FRQS) and the Quebec universities. The views expressed in this paper are those of the authors and not necessarily those of the CRDCN or its partners.

References

(1) Groome PA, Schulze KM, Keller S, Mackillop WJ. Demographic Differences between Cancer Survivors and those who Die Quickly of their Disease. Clin Oncol. 2008;20(8):647-56. DOI: https://doi.org/10.1016/j.clon.2008.05.006

(2) Canadian Cancer Society's Advisory Committee. Canadian Cancer Statistics: A 2020 special report on lung cancer. Toronto, ON: Canadian Cancer Society; 2020.

(3) Canadian Partnership Against Cancer. Lung cancer and equity: A focus on income and geography. Toronto, ON.: Canadian Partnership Against Cancer; 2020.

(4) Qureshi S, Boily G, Boulanger J, Pagé É, Strumpf E. Inequalities in survival and care across social determinants of health in a cohort of advanced lung cancer patients in Quebec (Canada): A high-resolution population-level analysis. Cancer Med. 2023;12(11):12683-704. DOI: 10.1002/cam4.5897

(5) Ruco A, Lofters AK, Lu H, et al. Lung cancer survival by immigrant status: a population-based retrospective cohort study in Ontario, Canada. BMC Cancer. 2024;24(1):1114. DOI: 10.1186/s12885-024-12804-7

(6) Qureshi S, Schnitzer ME, Nandi A, Strumpf E. Social inequalities in survival among Canadian lung cancer patients and the mediating role of stage at diagnosis. 2024.

(7) Malagón T, Botting-Provost S, Moore A, El-Zein M, Franco EL. Inequalities in relative cancer survival by race, immigration status, income, and education for 22 cancer sites in Canada, a cohort study. medRxiv. 2024:2024.05.31.24307976. DOI: 10.1101/2024.05.31.24307976

(8) Solar O, Irwin A. A conceptual framework for action on the social determinants of health. WHO Document Production Services; 2010.

(9) World Health Organization. Closing the gap in a generation - Health equity through action on. Geneva, Switzerland: Atlantic Collaborative on Injury Prevention; 2008. Available from: https://canadacommons.ca/kbart/artifacts/1886481/closing-the-gap-in-a-generation/2635796.

(10) National Academies of Sciences E, Medicine. Immigration as a Social Determinant of Health: Proceedings of a Workshop. Olson S, Anderson KM, editors. Washington, DC: The National Academies Press; 2018. 76 p.

(11) Mikkonen J, Raphael D. Social Determinants of Health: A Quick Guide for Health Professionals. York Univ Sch Heal Policy Manag. 2013:1-10. DOI:

(12) Bourgois P, Holmes SM, Sue K, Quesada J. Structural Vulnerability: Operationalizing the Concept to Address Health Disparities in Clinical Care. Acad Med. 2017;92(3). DOI: https://journals.lww.com/academicmedicine/fulltext/2017/03000/structural_vulnerability__operationalizing_the.18.aspx

(13) Donkin A, Goldblatt P, Allen J, Nathanson V, Marmot M. Global action on the social determinants of health. BMJ Global Health. 2018;3(Suppl 1):e000603. DOI: 10.1136/bmjgh-2017-000603

(14) Hankivsky O, Jordan-Zachery JS. The Palgrave Handbook of Intersectionality in Public Policy: Springer International Publishing; 2019. Available from: https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5667407.

(15) Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer survival: a review. Ann Oncol. 2006;17(1):5-19. DOI: https://doi.org/10.1093/annonc/mdj007

(16) Ahmed S, Shahid RK. Disparity in cancer care: a Canadian perspective. Curr Oncol. 2012;19(6):e376-e82. DOI: 10.3747/co.19.1177

(17) Maddison AR, Asada Y, Urquhart R. Inequity in access to cancer care: a review of the Canadian literature. Cancer Causes Control. 2011;22(3):359-66. DOI: 10.1007/s10552-010-9722-3

(18) Gillan C, Briggs K, Goytisolo Pazos A, et al. Barriers to accessing radiation therapy in Canada: a systematic review. Radiation Oncology. 2012;7(1):167. DOI: 10.1186/1748-717X-7-167

(19) Younis T, Al-Fayea T, Virik K, Morzycki W, Saint-Jacques N. Adjuvant Chemotherapy Uptake in Non-small Cell Lung Cancer. J Thorac Oncol. 2008;3(11):1272-8. DOI: https://doi.org/10.1097/JTO.0b013e318189f562

(20) Evans WK, Stiff J, Woltman KJ, et al. How equitable is access to treatment for lung cancer patients? A population-based review of treatment practices in Ontario. Lung cancer management. 2017;6(3):77-86. DOI: 10.2217/lmt-2017-0013

(21) Akhtar-Danesh GG, Finley C, Seow HY, Shakeel S, Akhtar-Danesh N. Change in treatment modality and trends in survival among stage I non-small cell lung cancer patients: a population-based study. J Thorac Dis. 2020;12(9):4670-9. DOI: 10.21037/jtd-20-1387

(22) Cook S, Alibhai S, Mehta R, et al. Improving Care for Older Adults with Cancer in Canada: A Call to Action. Current Oncology. 2024;31(7):3783-97. DOI: https://www.mdpi.com/1718-7729/31/7/279

(23) Sigel K, Bonomi M, Packer S, Wisnivesky J. Effect of Age on Survival of Clinical Stage I Non-Small-Cell Lung Cancer. Ann Surg Oncol. 2009;16(7):1912-7. DOI: 10.1245/s10434-009-0475-8

(24) Maccormick RE, Mackinnon KJ. Decision making in cancer treatment: age and socio-economic status as independent variables. Med Teach. 1990;12(3-4):353-5. DOI: 10.3109/01421599009006642

(25) Haase KR, Sattar S, Pilleron S, et al. A scoping review of ageism towards older adults in cancer care. J Geriatr Oncol. 2023;14(1):101385. DOI: https://doi.org/10.1016/j.jgo.2022.09.014

(26) Lofters AK, Gatov E, Lu H, et al. Lung Cancer Inequalities in Stage of Diagnosis in Ontario, Canada. Current Oncology. 2021;28(3):1946-56. DOI: https://doi.org/10.3390%2Fcurroncol28030181

(27) Ellison LF. Differences in cancer survival in Canada by sex. Health Rep. 2016;27(4):19-27. DOI:

(28) Cook MB, McGlynn KA, Devesa SS, Freedman ND, Anderson WF. Sex disparities in cancer mortality and survival. Cancer Epidemiol Biomarkers Prev. 2011;20(8):1629-37. DOI: 10.1158/1055-9965.Epi-11-0246

(29) Zafar A, Baessler F, Ihrig A, et al. Barriers to access cancer-related services for men in high-income countries: A narrative review looking beyond socioeconomic disadvantages. Aging and Cancer. 2022;3(3-4):147-60. DOI: https://doi.org/10.1002/aac2.12059

(30) Miller PA, Sinding C, McGillicuddy P, et al. Disparities in cancer care: Perspectives from the front line. Palliative and Supportive Care. 2014;12(3):175-81. DOI: 10.1017/S147895151200106X

(31) Varcoe C, Habib S, Sinding C, Fitch M, Magee Gullatte M, Cummings GG. Health disparities in cancer care: Exploring Canadian, American and international perspectives. 2015. 2015;25(1):9. DOI: DOI: 10.5737/236880762517381

(32) McDonald JT, Kennedy S. Cervical Cancer Screening by Immigrant and Minority Women in Canada. Journal of Immigrant and Minority Health. 2007;9(4):323-34. DOI: 10.1007/s10903-007-9046-x

(33) Canadian Partnership Against Cancer. Examining disparities in cancer control: A system performance special focus report. Canadian Partnership Against Cancer; 2014.

(34) Canadian Partnership Against Cancer. Cancer System Performance. 2017 Report. Toronto, ON.: Canadian Partnership Against Cancer; 2017.

(35) Srugo SA, Jiang Y, de Groh M. At-a-glance-Living arrangements and health status of seniors in the 2018 Canadian Community Health Survey. Health Promotion and Chronic Disease Prevention in Canada: Research, Policy and Practice. 2020;40(1):18. DOI: https://doi.org/10.24095/hpcdp.40.1.03

(36) Sinding C, Wiernikowski J. Treatment Decision Making and Its Discontents. Soc Work Health Care. 2009;48(6):614-34. DOI: 10.1080/00981380902831303

(37) Puts MT, Monette J, Girre V, et al. Characteristics of older newly diagnosed cancer patients refusing cancer treatments. Support Care Cancer. 2010;18(8):969-74. DOI: 10.1007/s00520-010-0883-0

(38) Pilleron S, Gower H, Janssen-Heijnen M, et al. Patterns of age disparities in colon and lung cancer survival: a systematic narrative literature review. BMJ Open. 2021;11(3):e044239. DOI: 10.1136/bmjopen-2020-044239

(39) Lofters AK, Hwang SW, Moineddin R, Glazier RH. Cervical cancer screening among urban immigrants by region of origin: A population-based cohort study. Prev Med. 2010;51(6):509-16. DOI: https://doi.org/10.1016/j.ypmed.2010.09.014

(40) Habib SZ. South Asian immigrant women's access to and experiences with breast and cervical cancer screening services in Canada [Text]2012.

(41) Lambert LK, Horrill TC, Beck SM, et al. Health and healthcare equity within the Canadian cancer care sector: a rapid scoping review. Int J Equity Health. 2023;22(1):20. DOI: 10.1186/s12939-023-01829-2

(42) 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. International Journal for Equity in Health. 2023;22(1):58. DOI: 10.1186/s12939-023-01860-3

(43) Horrill TC, Browne AJ, Stajduhar KI. Equity-Oriented Healthcare: What It Is and Why We Need It in Oncology. Curr Oncol. 2022;29(1):186-92. DOI: 10.3390/curroncol29010018

(44) Sayani A. Socially based inequities in breast cancer care: intersections of the social determinants of health and the cancer care continuum. Critical Studies: An International and Interdisciplinary Journal. 2017;13(1):24-36. DOI: https://doi.org/10.51357/cs.v13i1.129

(45) Postal Code OM Conversion File Plus (PCCF+) [Internet]. Government of Canada. 2023 [cited 2023-12-06]. Available from: https://www150.statcan.gc.ca/n1/en/catalogue/82F0086X#wb-auto-2.

(46) Akhtar-Danesh N, Finley C. Temporal trends in the incidence and relative survival of non-small cell lung cancer in Canada: A population-based study. Lung Cancer. 2015;90(1):8-14. DOI: https://doi.org/10.1016/j.lungcan.2015.07.004

(47) Ellison LF, Saint-Jacques N. Five-year cancer survival by stage at diagnosis in Canada. Health Rep. 2023;34(1):3-15. DOI: https://doi.org/10.25318/82-003-x202300100001-eng

(48) Ellis PM, Verma S, Sehdev S, Younus J, Leighl NB. Challenges to Implementation of an Epidermal Growth Factor Receptor Testing Strategy for Non–Small-Cell Lung Cancer in a Publicly Funded Health Care System. J Thorac Oncol. 2013;8(9):1136-41. DOI: https://doi.org/10.1097/JTO.0b013e31829f6a43

(49) 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;17(6):1471-4. DOI: https://doi.org/10.1245/s10434-010-0985-4

(50) Gasparini A. comorbidity: An R package for computing comorbidity scores. Journal of Open Source Software. 2018;3(23):648. DOI:

(51) R Core Team. R: A Language and Environment for Statistical Computing. 2023. Vienna, Austria. Available from: https://www.R-project.org/.

(52) Evans CR, Leckie G, Subramanian SV, Bell A, Merlo J. A tutorial for conducting intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA). SSM - Population Health. 2024;26:101664. DOI: https://doi.org/10.1016/j.ssmph.2024.101664

(53) Devine S, Uanhoro JO, Otto AR, Flake JK. Approaches for quantifying the ICC in multilevel logistic models: a didactic demonstration. Collabra: Psychology. 2024;10(1). DOI: https://doi.org/10.1525/collabra.94263

(54) Hox JJ, Moerbeek M, Schoot Rvd. Multilevel analysis : techniques and applications. New York: Routledge; 2018. Available from: https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1594541.

(55) Physicians for a Smoke-Free Canada. Tobacco Use 2000-2014: Insights from The Canadian Community Health Survey. 2016.

(56) Isaranuwatchai W, de Oliveira C, Mittmann N, et al. Impact of smoking on health system costs among cancer patients in a retrospective cohort study in Ontario, Canada. BMJ Open. 2019;9(6). DOI: 10.1136/bmjopen-2018-026022

(57) Canadian Cancer Society's Advisory Committee in collaboration with the Canadian Cancer Society, Statistics Canada and the Public Health Agency of Canada. Canadian Cancer Statistics 2023. Toronto, ON: Canadian Cancer Society; 2023.

(58) Khuder SA. Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. Lung Cancer. 2001;31(2-3):139-48. DOI: 10.1016/s0169-5002(00)00181-1

(59) Singh Setia M, Lynch J, Abrahamowicz M, Tousignant P, Quesnel-Vallee A. Self-rated health in Canadian immigrants: Analysis of the Longitudinal Survey of Immigrants to Canada. Health & Place. 2011;17(2):658-70. DOI: https://doi.org/10.1016/j.healthplace.2011.01.006

(60) Vang Z, Sigouin J, Flenon A, Gagnon A. The healthy immigrant effect in Canada: A systematic review. Population Change and Lifecourse Strategic Knowledge Cluster Discussion Paper Series/Un Réseau stratégique de connaissances Changements de population et parcours de vie Document de travail. 2015;3(1):4. DOI:

(61) Guruge S, Birpreet B, Samuels-Dennis JA. Health Status and Health Determinants of Older Immigrant Women in Canada: A Scoping Review. J Aging Res. 2015;2015:393761. DOI: 10.1155/2015/393761

(62) Bacal V, Blinder H, Momoli F, Wu KY, McFaul S. Is Immigrant Status Associated With Cervical Cancer Screening Among Women in Canada? Results From a Cross-Sectional Study. Journal of obstetrics and gynaecology Canada : JOGC = Journal d'obstetrique et gynecologie du Canada : JOGC. 2019;41(6):824-31.e1. DOI: 10.1016/j.jogc.2018.07.010

(63) Kobayashi KM, Prus SG. Examining the gender, ethnicity, and age dimensions of the healthy immigrant effect: factors in the development of equitable health policy. Int J Equity Health. 2012;11:8. DOI: 10.1186/1475-9276-11-8

(64) Ickert C, Senthilselvan A, Jhangri GS. Multilevel Modeling of Health Inequalities at the Intersection of Multiple Social Identities in Canada. The Sociological Quarterly. 2022;63(2):214-46. DOI: 10.1080/00380253.2020.1868956

(65) Booth CM, Li G, Zhang-Salomons J, Mackillop WJ. The impact of socioeconomic status on stage of cancer at diagnosis and survival. Cancer. 2010;116(17):4160-7. DOI: https://doi.org/10.1002/cncr.25427

(66) Crossman E. Executive Summary. Age composition of immigrant groups within the labour force survey. In: Canada CaI, editor.: Government of Canada; 2012.

(67) Statistics Canada. 2016 Census of Population: Immigration and ethnocultural diversity. In: Canada Go, editor.: Minister of IIndustry; 2017.

(68) Hou F, Schimmele C, Stick M. Changing demographics of racialized people in Canada. Economic and Social Reports [Internet]. 2023; 3(8):[18 p.]. Available from: http://central.bac-lac.gc.ca/.redirect?app=damspub&id=b5d8744a-4c8d-48f2-8a8f-3f58add37824.

(69) O'Grady JR, Ferdus J, Leylachian S, et al. Lung cancer in First Nations, Inuit, and Métis peoples in Canada - a scoping review. Int J Circumpolar Health. 2024;83(1):2381879. DOI: 10.1080/22423982.2024.2381879

(70) Evans CR, Leckie G, Merlo J. Multilevel versus single-level regression for the analysis of multilevel information: The case of quantitative intersectional analysis. Soc Sci Med. 2020;245:112499. DOI: https://doi.org/10.1016/j.socscimed.2019.112499

(71) Bauer GR, Scheim AI. Methods for analytic intercategorical intersectionality in quantitative research: Discrimination as a mediator of health inequalities. Soc Sci Med. 2019;226:236-45. DOI: https://doi.org/10.1016/j.socscimed.2018.12.015

(72) Evans CR, Williams DR, Onnela J-P, Subramanian SV. A multilevel approach to modeling health inequalities at the intersection of multiple social identities. Soc Sci Med. 2018;203:64-73. DOI: https://doi.org/10.1016/j.socscimed.2017.11.011

6.3. Supplementary materials

6.3.1. Appendix A

Databases used and information retrieved

Table 6.3.1. Data sources

Databases	Linkage years [A] (years used in the study)	Information extracted
CCR TMF	1992-2015 (2002-2013)	age sex date of lung cancer diagnosis cancer stage tumour histology postal code date of death [B]
CVSD	1992-2014 (2002-2014)	date of death [B]
T1PMF	1992-2014 (2002-2014)	date of death [B]
IMDB	1980-2015 (1980-2013)	immigrant status
T1FF [C]	1992-2015 (1999-2012)	income
DAD	1994/95-2015/16 (2004/05-2013/14)	comorbidities
NACRS	2002/03-2015/16 (2004/05-2013/14)	comorbidities
PCCF+ [D]	Census (SAS program versions): census 2001 (4k) census 2006 (5k) census 2011 (6d)	province Statistical Area Classification type

[A] Although the linked data covered many years, only information relevant for patients diagnosed between 2002 and 2013 was used in our study.

[B] Information on death in the CCR was supplemented with death data from CVSD and T1PMF files.

[C] Records from the T1FF which corresponded to the year in which individuals were diagnosed with lung cancer diagnosis were often incomplete, likely due to individuals dying within the same year.

[D] PCCF+ version 4k used for postal codes in diagnosis years 2002 to 2004, version 5k for postal codes in diagnosis years 2005 to 2009 and 6d was used for postal codes in diagnosis years 2010 to 2013.

Abbreviations: CCR TMF: Canadian Cancer Registry incidence Tabulation Master File; IMDB: Longitudinal Immigration Database; T1FF: T1 Family File; CVSD: Canadian Vital Statistics Database; T1PMF: T1 Personal Master File; DAD: Discharge Abstract Database; NACRS: National Ambulatory Care Reporting System; and PCCF+: Postal Code Conversion File Plus

1. Cancer and Death Data from the Canadian Cancer Registry (CCR) linked to Canadian Vital Statistics Database (CVSD) and the T1 Personal Master File (T1PMF)

Statistics Canada's Canadian Cancer Registry (CCR) includes the CCR Tabulation Master File (CCR TMF), a dataset that contains detailed information on all cancers diagnosed in Canada since 1992. The data in the CCR TMF comes from cancer records submitted to the CCR by provinces and territories from their own registries. The information submitted is standardized to facilitate the identification of multiple primary tumours.

Death-linked analytical file for selection of lung cancer tumours

Using the CCR TMF with diagnoses from January 1, 1992, to December 31, 2015, Statistics Canada created a cohort for mortality follow-up until December 31, 2014. With the objective of ascertaining vital status, this cohort was linked to two databases that also include death information: the Canadian Vital Statistics Database (CVSD, years 1992-2014) and the T1 Personal Master File (T1PMF, years 1992-2014). The death-linked file was further processed by Statistics Canada to create the death-linked *analytical* file in which multiple primary tumours are identified according to the International Agency for Research on Cancer (IARC) rules.[1]

We identified all primary diagnoses of lung and bronchus (lung) cancers in the death-linked analytical file using CCR's adapted version of the Surveillance, Epidemiology, and End Results grouping methodology (SEER code 0302).[2] The SEER group for lung cancers includes all cancers with topography code C34 and histology codes 8000 to 9049, 9056 to 9139, and 9141 to 9589 according to the International Classification of Diseases for Oncology, third edition (ICD-O-3).[3] Lung cancers reported between 1992 and 2006 were also identified using the International Classification of Diseases, ninth revision, diagnostic codes (ICD-9 codes 162.2 to 162.9).[4] For these cancers, we excluded histology codes for mesotheliomas (9050–9055), Kaposi sarcomas (9140), or other hematopoietic and lymphoid neoplasms (9590–9992) codes according to the International Classification of Diseases for Oncology, second edition (ICD-O-2).[5]

Removal of multiple lung cancer primary tumours and selection of cancers between 2002 and 2013

There were 2,830 out of 464,370 individuals (0.6%) in the death-linked analytical file with multiple primary lung cancer diagnoses. We applied the following rules to keep only one primary lung cancer diagnosis per individual.

1. If dates of diagnosis and stages of cancer are identical: suppress any duplicate record

2. If dates of diagnosis are identical and stages of cancer are different: suppress record with lower stage or with stage missing

3. If dates of diagnosis are different and beyond 30 days apart: suppress most recent record

4. If dates of diagnosis are different and within 30 days apart, and stages of cancer are identical: suppress the most recent record

5. If dates of diagnosis are different and within 30 days apart and stages of cancer are different: suppress record with lower stage or with stage missing

After restricting the death-linked analytical file to one primary lung cancer diagnosis per individual, we selected all individuals diagnosed between 2002 and 2013 for our study. This period allowed us to gather the information necessary to conduct our analyses, such as income two years prior to cancer diagnosis (see below), and adequate one-year follow-up for vital status (i.e., 1-year survival status).

Defining lung cancer histologic type

Based on the Canadian Cancer Society's Advisory Committee's histological groupings, lung cancer ICD-O-3 histology codes were combined into the following seven groups: squamous cell carcinoma, adenocarcinoma, small cell carcinoma, large cell carcinoma, non-small cell lung cancer (not otherwise specified), sarcomas and other specified neoplasms, and unspecified.[6]

The Canadian Cancer Society's Advisory Committee's histological groupings	
Histologic type	**ICD-O-3 Histology code**
Squamous cell carcinoma	8050-8078
Adenocarcinoma	8140, 8211, 8230, 8231, 8250-8260, 8323, 8480-8490,8550-8552, 8570-8574,8576
Small cell carcinoma	8040-8045
Large cell carcinoma	8011-8012,8014-8031, 8035,8310
Non-small cell lung cancer, not otherwise specified (NOS)	Excluding codes above: 8010-8576
Sarcomas and other specified neoplasms	8800-8811, 8830, 8840-8921, 8990-8991, 9040-9044, 9120-9133, 9150, 9540-9581, 8815, 8980
Unspecified	8000-8005

Death Information

Statistics Canada linked the CCR TMF to the Canadian Vital Statistics Database (CVSD) and the T1 Personal Master File (T1PMF) to supplement it with death-related information, such as the date and cause of death.

The CVSD is a Statistics Canada database that compiles information on all deaths in Canada annually through events collected and reported by each of the provincial and territorial vital statistics registries. In Canada, all individuals who earn taxable income are required to file an income tax return (i.e., T1 form) on an annual basis. The T1PMF is a file compiled by the Canada Revenue Agency (CRA) and provided to Statistics Canada, which combines data on all income tax filers or those who received the Canada Child Tax Benefit (CCTB), including their children and spouses. The T1PMF-D, a subset of the T1PMF that contains only death information, is used to supplement the CCR TMF.

Statistics Canada primarily derived the date of death in the death-linked analytical file using death information in the CVSD. Death information from the T1PMF-D was only used to validate dates when there were discrepancies between the CCR and the CVSD or to account for deaths not found in the CVSD. Death information from the CCR was only used when it was missing from the CVSD and the T1PMF-D.

2. Income Data from the Annual Income Estimates for Census Families and Individuals file (T1 Family File or T1FF)

The T1FF is another file compiled by the CRA and provided to Statistics Canada. It contains data from income tax returns and the CCTB. The T1FF includes information on sources of income and demographic indicators derived from data on tax filers and non-tax filers identified in income tax returns or the CCTB.

Individuals with lung cancer in the death-linked *analytical* file were linked to T1FF files for taxation years 1992 to 2014. These files included the family after-tax income (excluding capital gains/losses), which was divided by the square root of the family size to account for economies of scale in family expenditures (i.e., adjusted or equivalized family after-tax income).[7] For our study, we referred to the adjusted family after-tax income reported in T1FF records for the taxation year corresponding to two years prior to individuals' lung cancer diagnosis. We opted not to use

information from T1FF from the taxation year corresponding to the year of a lung cancer diagnosis due to a high proportion of missing data (i.e., low linkage rates) at that time, which is likely caused by respondents dying within the same year. We also did not use income in the taxation year corresponding to one year prior to a lung cancer diagnosis (i.e., the period shortly before diagnosis) due to potential changes in income from a decline in health.

Subsequent to imputations for missing adjusted family after-tax income or missing family-level place of residence at the time of the reported income, we ranked our income measure into population-level terciles of family income referenced by calendar year, province, and place of residence (urban-rural area: large urban cities, small urban cities, rural areas, and territories). See below. Only income terciles assigned to T1FF observations in the taxation years 1999 to 2012 were relevant for our study (i.e., two years prior to cancer diagnoses in 2002 to 2013 and accounting for years used for imputation - see below).

Income imputation

Missing income information two years before diagnosis was imputed by taking the average of non-missing incomes one and three years before diagnosis. If income was missing one year before diagnosis, then the non-missing income at three years was carried forward to impute the missing income at two years. In contrast, the imputation of missing income two years before diagnosis was not carried out when income was also missing three years before diagnosis, or if it was missing both at three years and one year before diagnosis.

Place of residence at time of reported income and imputation

Statistics Canada used PPCF+ files to code postal codes in the T1FF files into census-defined geographic areas. Each family income reported in the T1FF was assigned a census-specific geographic coordinate based on the family-level postal code (CMAs: census metropolitan areas that are large urban cities; CAs: census agglomerations that are small urban cities; non-CMA/CA rural areas; and non-CMA/CA areas that are territories) as follows:

T1FF family income from taxation year:	Family CMA, CA, and non-CMA/CA according to census year:
1992	1986
1993-1995	1991
1996-1999	1996
2000-2004	2001
2005-2009	2006
2010-2014	2011

Using T1FF data from three years prior to individuals' diagnosis, we imputed geographic coordinates when they were missing two years before diagnosis. The non-missing geographic coordinate in the year before the missing value was first carried forward, and then the non-missing geographic coordinate in the year after was carried backward for the remaining missing values. If geographic coordinates were also missing in the year before and after the missing value two years prior to diagnosis, then no imputation was carried out.

Population-level income terciles

We used population-level income terciles referenced by calendar year, province, and place of residence (urban-rural area: large urban cities, small urban cities, rural areas, and territories) that were created in our previous work.[8] Briefly, respondents of the Canadian Community Health Surveys conducted between the years 2000 and 2017 were linked to T1FF data. The data on adjusted family after-tax income was then divided into terciles according to the survey year, province, and place of residence based on self-reported postal codes. Survey weights were taken into consideration to create terciles at the level of the Canadian population, excluding the Canadian Armed Forces, and those who are institutionalized or living on reserves.

3. Immigration Data from the Longitudinal Immigration Database (IMDB)

The IMDB includes data on immigrant admissions and permits released for non-permanent residency in Canada, obtained from Immigration, Refugees and Citizenship Canada (IRCC). The IMDB includes individual-level sociodemographic and immigration data on immigrants and non-permanent residents who landed in Canada between 1980 and 2014. Individuals who landed in Canada before this period are not identifiable in the IMDB. For our study, all individuals diagnosed with a primary lung cancer between 2002 and 2013 who were associated with an IMDB ID in an IMDB linkage key file were identified as immigrants.

4. Comorbidity Data from the Discharge Abstract Database (DAD) and the National Ambulatory Care Reporting System (NACRS)

DAD is an administrative database that includes data on hospital discharges from acute care for individuals in all Canadian provinces and territories, except Quebec. Data from the fiscal years 1994/95 to 2014/15 was linked to the CCR: DAD-CCR.

NACRS is an administrative database that includes information on visits related to ambulatory care in hospital and community-based settings (e.g., outpatient clinics), such as day surgeries, emergency department visits, diagnostic imaging, and oncology. Data from participating jurisdictions in NACRS from fiscal years 2002/03 to 2015/16 was linked to the CCR: DAD-NACRS.

It is important to note that the data reported in DAD and NACRS varies with time and jurisdiction. For example, DAD has integrated information on long-term care over time, and the reporting of emergency department visits in NACRS is not mandated for Newfoundland and Labrador, New Brunswick, the Northwest Territories, and Nunavut.

Data on comorbidities

To create a comorbidity index, we first extracted all records from the DAD-CCR and the NACRS-CCR datasets pertaining to individuals with a primary diagnosis of lung cancer between 2010 and 2013, who were included in our study's secondary analyses. In each dataset, we combined all diagnostic codes associated with one individual within five years prior to the date of lung cancer diagnosis. This meant going all the way back to data from 2005 (i.e., fiscal year 2004/05) for patients diagnosed in 2010. All diagnosis codes reported in DAD and NACRS were coded with the International Classification of Diseases, 10th Revision, starting in fiscal years 2004/05 and 2002/03, respectively.

In DAD-CCR, we combined data from up to 25 diagnostic code variables, each representing a condition related to a patient's hospitalization, and aggregated data from multiple observations per ID (i.e., multiple events per hospitalization and/or multiple hospitalizations). Similarly, in NACRS-CCR, we combined data from up to 10 diagnostic code variables, each representing a condition related to a patient's visit, and aggregated data from multiple observations per ID (i.e., multiple events per visit and/or multiple visits). The aggregated information from DAD-CCR and

NACRS-CCR was then combined into one dataset, after which all ICD-10 codes related to lung cancer (i.e., diagnosis codes starting with 162, or starting with C34) were removed from the aggregated data.

The *comorbidity* package in R was then used on the combined DAD and NACR data to obtain a weighted Charlson index (coded as 0, 1-2, 3-4, or ≥5 comorbidities).[9-11]

5. Geographical Data from the Postal Code Conversion Files – Plus (PCCF+)

Each individual's 6-digit postal code can be linked to standard geographic areas that have been defined by Statistics Canada based on the population census.[12] The linkage is carried out through the PCCF+ program developed in the statistical software SAS. Since one postal code can link to multiple standard geographic areas, each postal code is assigned a unique geographic area in a probabilistic manner according to population weights derived from the census. The PCCF (not PCCF+) file, which links each postal code to a unique geographic area in a constant manner by selecting the area with the most dwellings, was not used because it is less accurate than the PCCF+ in assigning rural postal codes.[13]

Creation of urban-rural spectrum variable

The standard geographic area defined as Statistical Area Classification (SAC) assigned to each postal code was used to create the urban-rural exposure. SAC types group small geographic areas defined as census subdivisions or municipalities based on population size, density, and commuting flows into Census Metropolitan Areas (CMAs), Census Agglomerations (CAs), Metropolitan Influenced Zones (MIZs), and territories. CMAs have a population of 100,000+, of which 50,000+ reside in the core; CAs have a core population of 10,000+; MIZs consist of municipalities outside of CMAs/CAs with strong, moderate, weak, or no metropolitan influence; and territories include all areas in the territories that are not a CA (i.e., Yellowknife).

Each individual's postal code reported in the CCR at the time of diagnosis was entered into the PCCF+ program in SAS (PCCF+ versions for postal codes in years 2002-2004: 4k, 2005-2009: 5k, and 2010-2013: 6d) to determine the SAC type of their place of residence. If the postal code at the time of diagnosis was missing, it was imputed using non-missing postal codes in the TIFF file with the taxation year corresponding to the diagnosis year. For the urban-rural spectrum exposure,

we categorized SAC types into 1) large urban cities (CMAs), 2) small urban cities (CAs), and 3) rural areas (all MIZs and territories).

6.3.2. Appendix B

Main analysis: strata frequencies, predicted probabilties, and deviations

Table 6.3.2. Main analysis: strata frequencies, predicted probabilities, and intersectional effects
Study period: 2002-2013, N= 183,350.

Stratum ID	Age	Sex	Income tercile	Immigrant status	N	Predicted 1-year survival probability in % (95% CI) - model 1B	Difference in predicted from expected 1-year survival probability in % (95% CI) - model 1B
1111	< 60	Male	1: High	Long-term resident	4,040	45.24 (43.23, 47.38)	-2.00 (-3.98, -0.02)
1112	< 60	Male	1: High	Immigrant	255	59.71 (56.80, 62.53)	1.03 (-2.13, 4.20)
1121	< 60	Male	2: Medium	Long-term resident	4,055	42.61 (40.63, 44.51)	-0.99 (-3.02, 1.03)
1122	< 60	Male	2: Medium	Immigrant	505	54.50 (51.66, 57.46)	-0.61 (-3.76, 2.54)
1131	< 60	Male	3: Low	Long-term resident	5,405	36.31 (34.44, 38.19)	-2.78 (-5.01, -0.54)
1132	< 60	Male	3: Low	Immigrant	735	51.55 (48.73, 54.11)	1.05 (-1.70, 3.80)
1211	< 60	Female	1: High	Long-term resident	4,590	57.43 (55.38, 59.19)	1.18 (-0.69, 3.05)
1212	< 60	Female	1: High	Immigrant	255	66.91 (63.93, 69.58)	-0.31 (-3.14, 2.53)
1221	< 60	Female	2: Medium	Long-term resident	4,855	53.38 (51.55, 55.26)	0.87 (-1.15, 2.89)
1222	< 60	Female	2: Medium	Immigrant	350	64.15 (61.25, 67.09)	0.27 (-2.71, 3.25)
1231	< 60	Female	3: Low	Long-term resident	6,165	46.89 (44.85, 48.82)	-1.13 (-3.10, 0.84)
1232	< 60	Female	3: Low	Immigrant	650	62.64 (59.85, 65.33)	3.02 (0.26, 5.78)
2111	60-64	Male	1: High	Long-term resident	2,980	46.06 (43.92, 48.26)	0.57 (-1.54, 2.69)
2112	60-64	Male	1: High	Immigrant	125	57.35 (53.96, 60.70)	0.12 (-3.10, 3.34)
2121	60-64	Male	2: Medium	Long-term resident	3,580	40.37 (38.38, 42.45)	-1.56 (-3.89, 0.76)
2122	60-64	Male	2: Medium	Immigrant	185	53.24 (49.83, 56.33)	-0.13 (-3.08, 2.83)
2131	60-64	Male	3: Low	Long-term resident	4,830	36.32 (34.44, 38.16)	-1.30 (-3.28, 0.69)
2132	60-64	Male	3: Low	Immigrant	320	50.02 (46.96, 53.10)	1.15 (-1.86, 4.17)
2211	60-64	Female	1: High	Long-term resident	2,295	53.64 (51.28, 55.93)	-0.92 (-3.12, 1.28)
2212	60-64	Female	1: High	Immigrant	70	65.51 (62.37, 68.55)	-0.18 (-3.09, 2.72)
2221	60-64	Female	2: Medium	Long-term resident	3,385	51.96 (49.86, 54.12)	1.04 (-1.02, 3.10)
2222	60-64	Female	2: Medium	Immigrant	95	62.49 (59.42, 65.49)	0.13 (-3.09, 3.34)
2231	60-64	Female	3: Low	Long-term resident	4,690	46.67 (44.62, 48.65)	0.36 (-1.82, 2.54)
2232	60-64	Female	3: Low	Immigrant	225	58.59 (55.46, 61.63)	0.74 (-2.14, 3.63)
3111	65-69	Male	1: High	Long-term resident	2,880	42.57 (40.20, 44.80)	-1.07 (-3.01, 0.86)

Table 6.3.2 (continued). Main analysis: strata frequencies, predicted probabilities, and intersectional effects
Study period: 2002-2013, N= 183,350.

Stratum ID	Age	Sex	Income tercile	Immigrant status	N	Predicted 1-year survival probability in % (95% CI) - model 1B	Difference in predicted from expected 1-year survival probability in % (95% CI) - model 1B
3112	65-69	Male	1: High	Immigrant	70	55.73 (52.01, 59.23)	0.61 (-2.86, 4.07)
3121	65-69	Male	2: Medium	Long-term resident	5,235	40.95 (38.99, 42.92)	0.90 (-1.25, 3.04)
3122	65-69	Male	2: Medium	Immigrant	125	50.68 (47.34, 54.35)	-0.77 (-3.93, 2.39)
3131	65-69	Male	3: Low	Long-term resident	6,785	35.16 (33.27, 36.83)	-0.83 (-2.86, 1.21)
3132	65-69	Male	3: Low	Immigrant	460	46.86 (43.71, 49.83)	-0.13 (-3.02, 2.77)
3211	65-69	Female	1: High	Long-term resident	2,010	53.57 (51.37, 56.05)	0.86 (-1.69, 3.42)
3212	65-69	Female	1: High	Immigrant	35	63.60 (60.26, 66.95)	-0.34 (-3.86, 3.17)
3221	65-69	Female	2: Medium	Long-term resident	4,450	49.87 (47.89, 51.85)	0.73 (-1.62, 3.07)
3222	65-69	Female	2: Medium	Immigrant	60	60.25 (56.77, 63.65)	-0.20 (-3.54, 3.14)
3231	65-69	Female	3: Low	Long-term resident	6,400	44.33 (42.43, 46.32)	-0.12 (-1.97, 1.74)
3232	65-69	Female	3: Low	Immigrant	295	56.77 (53.58, 60.02)	0.59 (-2.32, 3.51)
4111	70-74	Male	1: High	Long-term resident	2,600	42.60 (40.29, 44.76)	1.35 (-0.78, 3.48)
4112	70-74	Male	1: High	Immigrant	40	52.66 (49.22, 56.10)	-0.11 (-3.30, 3.09)
4121	70-74	Male	2: Medium	Long-term resident	6,440	37.99 (36.06, 39.77)	0.31 (-1.83, 2.46)
4122	70-74	Male	2: Medium	Immigrant	110	48.92 (45.54, 52.31)	-0.01 (-3.58, 3.55)
4131	70-74	Male	3: Low	Long-term resident	8,005	33.85 (32.06, 35.55)	0.34 (-1.54, 2.21)
4132	70-74	Male	3: Low	Immigrant	615	44.77 (41.88, 47.89)	0.31 (-2.64, 3.25)
4211	70-74	Female	1: High	Long-term resident	1,830	49.68 (47.24, 52.31)	-0.43 (-2.75, 1.89)
4212	70-74	Female	1: High	Immigrant	25	61.36 (57.91, 64.88)	-0.08 (-3.39, 3.22)
4221	70-74	Female	2: Medium	Long-term resident	4,720	46.33 (44.41, 48.42)	-0.20 (-2.02, 1.63)
4222	70-74	Female	2: Medium	Immigrant	45	57.69 (54.50, 61.13)	-0.23 (-3.85, 3.38)
4231	70-74	Female	3: Low	Long-term resident	7,260	40.82 (38.93, 42.64)	-1.16 (-3.34, 1.02)
4232	70-74	Female	3: Low	Immigrant	375	53.68 (50.52, 56.69)	0.19 (-2.82, 3.19)
5111	75-79	Male	1: High	Long-term resident	2,300	36.84 (34.91, 39.09)	0.32 (-2.17, 2.81)
5112	75-79	Male	1: High	Immigrant	35	47.43 (44.04, 50.82)	-0.50 (-3.74, 2.74)
5121	75-79	Male	2: Medium	Long-term resident	5,980	34.26 (32.43, 36.00)	1.19 (-0.94, 3.33)
5122	75-79	Male	2: Medium	Immigrant	75	43.22 (39.93, 47.14)	-0.83 (-4.12, 2.46)
5131	75-79	Male	3: Low	Long-term resident	7,905	30.40 (28.77, 32.15)	1.15 (-0.54, 2.85)

Table 6.3.2 (continued). Main analysis: strata frequencies, predicted probabilities, and intersectional effects
Study period: 2002-2013, N= 183,350.

Stratum ID	Age	Sex	Income tercile	Immigrant status	N	Predicted 1-year survival probability in % (95% CI) - model 1B	Difference in predicted from expected 1-year survival probability in % (95% CI) - model 1B
5132	75-79	Male	3: Low	Immigrant	585	39.45 (36.78, 42.21)	-0.15 (-3.05, 2.75)
5211	75-79	Female	1: High	Long-term resident	1,555	44.72 (42.39, 47.00)	-0.50 (-2.74, 1.74)
5212	75-79	Female	1: High	Immigrant	20	56.67 (53.04, 60.01)	-0.22 (-3.86, 3.41)
5221	75-79	Female	2: Medium	Long-term resident	4,080	41.00 (38.97, 42.93)	-0.59 (-2.78, 1.61)
5222	75-79	Female	2: Medium	Immigrant	40	52.91 (49.31, 56.26)	-0.30 (-3.70, 3.11)
5231	75-79	Female	3: Low	Long-term resident	7,630	37.20 (35.31, 39.06)	-0.14 (-2.21, 1.94)
5232	75-79	Female	3: Low	Immigrant	415	48.79 (45.92, 51.91)	0.25 (-2.69, 3.19)
6111	≥ 80	Male	1: High	Long-term resident	2,460	27.77 (25.89, 29.78)	0.39 (-1.74, 2.51)
6112	≥ 80	Male	1: High	Immigrant	45	37.30 (34.17, 40.55)	0.03 (-3.44, 3.50)
6121	≥ 80	Male	2: Medium	Long-term resident	6,555	25.59 (24.12, 27.17)	1.15 (-0.63, 2.93)
6122	≥ 80	Male	2: Medium	Immigrant	110	33.59 (30.58, 36.94)	-0.47 (-3.21, 2.27)
6131	≥ 80	Male	3: Low	Long-term resident	9,615	23.20 (21.91, 24.59)	1.90 (0.36, 3.43)
6132	≥ 80	Male	3: Low	Immigrant	845	29.81 (27.37, 32.12)	-0.39 (-2.99, 2.22)
6211	≥ 80	Female	1: High	Long-term resident	1,720	34.97 (32.70, 37.08)	-0.26 (-2.60, 2.09)
6212	≥ 80	Female	1: High	Immigrant	20	46.04 (42.45, 49.50)	-0.26 (-3.68, 3.16)
6221	≥ 80	Female	2: Medium	Long-term resident	4,435	32.24 (30.43, 34.28)	0.38 (-1.68, 2.44)
6222	≥ 80	Female	2: Medium	Immigrant	40	42.17 (38.99, 45.72)	-0.40 (-3.84, 3.05)
6231	≥ 80	Female	3: Low	Long-term resident	10,730	27.90 (26.51, 29.43)	-0.16 (-1.66, 1.34)
6232	≥ 80	Female	3: Low	Immigrant	665	35.33 (32.62, 38.02)	-2.84 (-5.40, -0.28)

6.3.3. Appendix C

Stratified analyses analyses: strata deviations

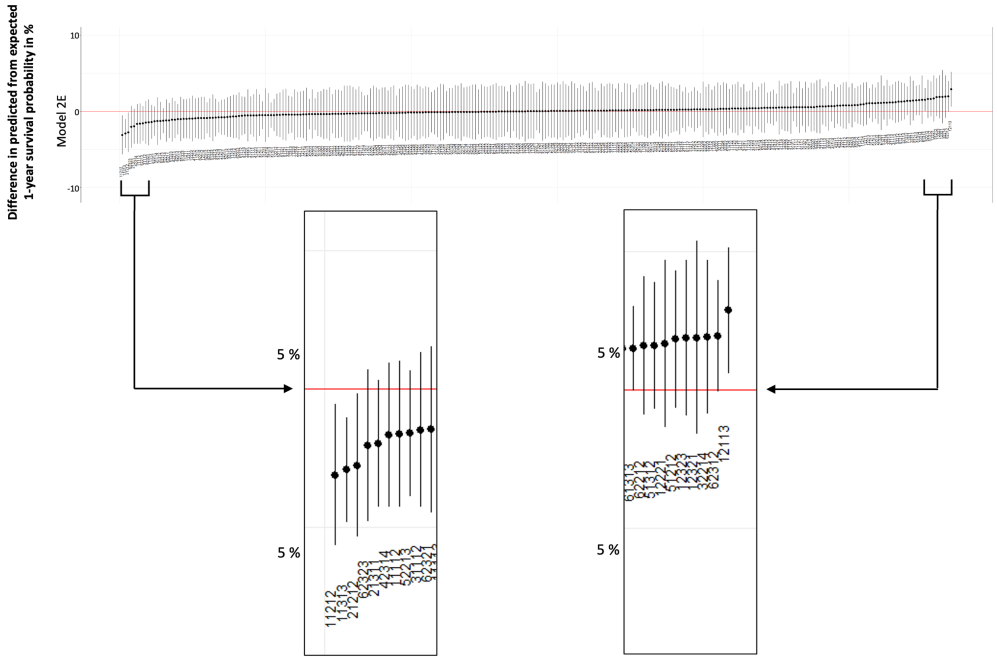


Figure 6.3.1. Stratified analysisis - province: stratum-specific differences in predicted 1-year survival probabilities from the expected probabilities based solely on fixed effects.

Study period: 2002-2013, N =183,350. Estimates obtained from main effects models 2E. The strata are ranked according to increasing differences, and the bottom and top 10 observations are presented in detail.

Strata ID defined with 5-digit codes with:

- the 1st position representing age in years (1: <60; 2: 60-64; 3: 65-69; 4: 70-74; 5: 75-79; and 6: ≥80)
- the 2nd position representing sex (1: male; and 2: female)
- the 3rd position representing family income tercile (1: high income, 2: medium income, and 2: low income)
- the 4th position representing immigration status (1: long-term resident; and 2: immigrant)
- the 5th digit for model 2E representing province category (1: BC; 2: MN and AB; 3: ON, NB, and NFL: 4: SK, NS, PEI, and YK/NWT/NU)

For example, the stratum "12311" in model 2E represents women aged <60 who are low-income, long-term residents of Canada, and live in BC.

Abbreviations: BC: British Columbia; MN: Manitoba; AB: Alberta; ON: Ontario, NB: New Brunswick, NFL: Newfoundland and Labrador; SK: Saskatchewan; NS: Nova Scotia; PEI: Prince Edward Island; YK: Yukon; NWT: Northwest Territories; NU: Nunavut.

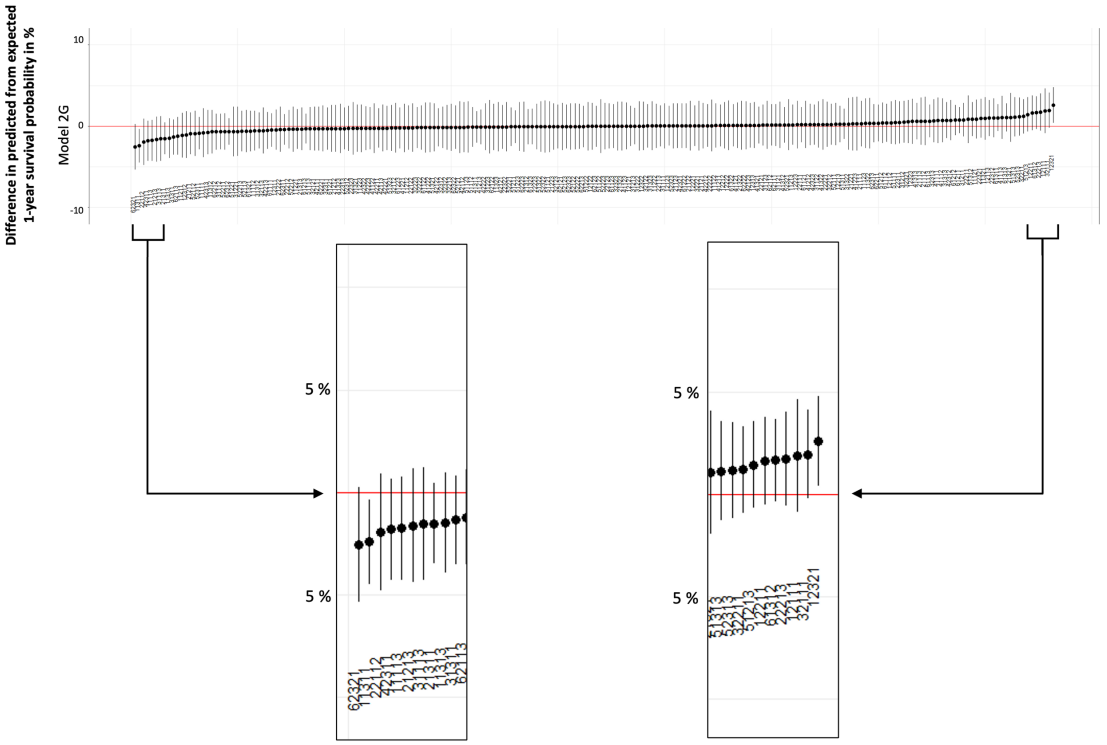


Figure 6.3.2. Stratified analysisis – urban/rural area: stratum-specific differences in predicted 1-year survival probabilities from the expected probabilities based solely on fixed effects.

Study period: 2002-2013, N =183,350. Estimates obtained from main effects models 2G. The strata are ranked according to increasing differences, and the bottom and top 10 observations are presented in detail.

Strata ID defined with 5-digit codes with:

- the 1st position representing age in years (1: <60; 2: 60-64; 3: 65-69; 4: 70-74; 5: 75-79; and 6: ≥80)
- the 2nd position representing sex (1: male; and 2: female)
- the 3rd position representing family income tercile (1: high income, 2: medium income, and 2: low income)
- the 4th position representing immigration status (1: long-term resident; and 2: immigrant)
- the 5th digit for model 2G representing urban-rural area (1: large urban cities, 2: small urban cities, 3: rural areas)

For example, the stratum "12311" in model 2G represents women aged <60 who are low-income, long-term residents of Canada, and live in large urban cities.

Abbreviations: BC: British Columbia; MN: Manitoba; AB: Alberta; ON: Ontario, NB: New Brunswick, NFL: Newfoundland and Labrador; SK: Saskatchewan; NS: Nova Scotia; PEI: Prince Edward Island; YK: Yukon; NWT: Northwest Territories; NU: Nunavut.

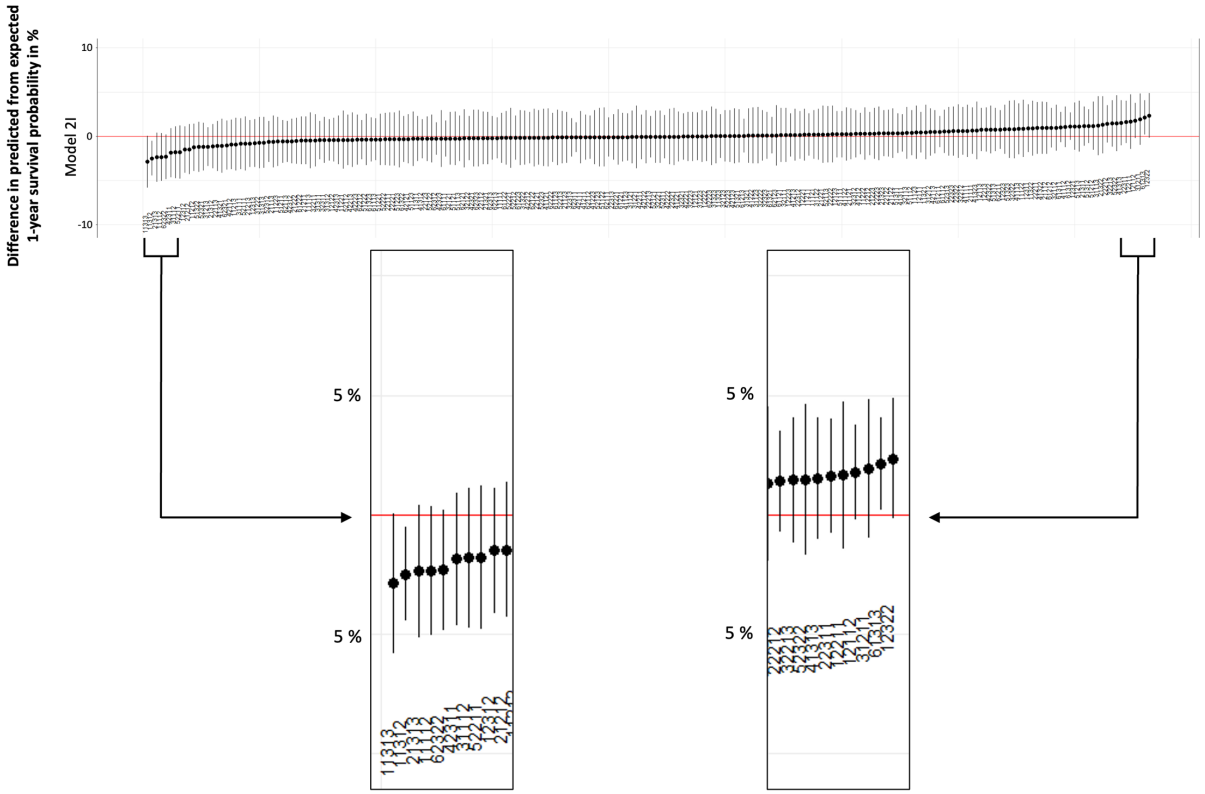


Figure 6.3.3. Stratified analysis – time period: stratum-specific differences in predicted 1-year survival probabilities from the expected probabilities based solely on fixed effects.

Study period: 2002-2013, N =183,350. Estimates obtained from main effects models 2I. The strata are ranked according to increasing differences, and the bottom and top 10 observations are presented in detail.

Strata ID defined with 5-digit codes with:

- the 1st position representing age in years (1: <60; 2: 60-64; 3: 65-69; 4: 70-74; 5: 75-79; and 6: ≥80)
- the 2nd position representing sex (1: male; and 2: female)
- the 3rd position representing family income tercile (1: high income, 2: medium income, and 2: low income)
- the 4th position representing immigration status (1: long-term resident; and 2: immigrant)
- the 5th digit for model 2I representing year of diagnosis (1: 2002-2004; 2: 2005-2010; 3: 2011-2013)

For example, the stratum "12311" in model 2G represents women aged <60 who are low-income, long-term residents of Canada, and live in large urban cities.

Abbreviations: BC: British Columbia; MN: Manitoba; AB: Alberta; ON: Ontario, NB: New Brunswick, NFL: Newfoundland and Labrador; SK: Saskatchewan; NS: Nova Scotia; PEI: Prince Edward Island; YK: Yukon; NWT: Northwest Territories; NU: Nunavut.

6.3.4. Appendix D

Secondary analysis: strata frequencies, predicted probabilties, and deviations

Table 6.3.3. Secondary analysis: strata frequencies, predicted probabilities, and intersectional effects
Study period: 2010-2013, N=54,735.

Strata ID	Age	Sex	Income tercile	Immigrant status	N	Predicted 1-year survival probability in % (95% CI) - model 3B	Difference in predicted from expected 1-year survival probability in % (95% CI) - model 3B
1111	< 60	Male	1: High	Long-term resident	1,045	49.17 (46.66, 51.68)	-0.53 (-3.17, 2.10)
1112	< 60	Male	1: High	Immigrant	100	61.54 (58.62, 64.50)	0.18 (-2.82, 3.17)
1121	< 60	Male	2: Medium	Long-term resident	1,180	46.01 (43.71, 48.52)	-0.03 (-2.44, 2.39)
1122	< 60	Male	2: Medium	Immigrant	205	57.39 (54.24, 60.34)	-0.64 (-3.59, 2.32)
1131	< 60	Male	3: Low	Long-term resident	1,505	40.43 (38.22, 42.87)	-1.34 (-3.92, 1.23)
1132	< 60	Male	3: Low	Immigrant	305	53.83 (50.66, 56.64)	0.22 (-2.58, 3.02)
1211	< 60	Female	1: High	Long-term resident	1,380	60.03 (57.61, 62.33)	1.03 (-1.34, 3.40)
1212	< 60	Female	1: High	Immigrant	105	70.02 (67.27, 72.69)	0.13 (-2.63, 2.90)
1221	< 60	Female	2: Medium	Long-term resident	1,500	56.03 (53.84, 58.27)	0.31 (-2.16, 2.78)
1222	< 60	Female	2: Medium	Immigrant	140	67.33 (64.37, 69.91)	0.38 (-2.41, 3.18)
1231	< 60	Female	3: Low	Long-term resident	1,885	50.23 (47.82, 52.57)	-1.01 (-3.15, 1.13)
1232	< 60	Female	3: Low	Immigrant	215	63.41 (60.48, 65.90)	0.58 (-2.11, 3.26)
2111	60-64	Male	1: High	Long-term resident	935	48.47 (45.94, 51.02)	0.08 (-2.52, 2.67)
2112	60-64	Male	1: High	Immigrant	60	60.06 (56.72, 62.96)	-0.04 (-3.29, 3.20)
2121	60-64	Male	2: Medium	Long-term resident	1,150	43.94 (41.55, 46.49)	-0.93 (-3.54, 1.67)
2122	60-64	Male	2: Medium	Immigrant	80	56.71 (53.43, 59.94)	0.05 (-3.20, 3.29)
2131	60-64	Male	3: Low	Long-term resident	1,435	39.58 (37.34, 41.81)	-0.75 (-3.24, 1.74)
2132	60-64	Male	3: Low	Immigrant	130	52.62 (49.50, 55.63)	0.45 (-2.57, 3.48)
2211	60-64	Female	1: High	Long-term resident	835	57.47 (54.96, 60.04)	-0.28 (-2.87, 2.31)
2212	60-64	Female	1: High	Immigrant	30	68.65 (65.64, 71.52)	-0.04 (-3.10, 3.02)
2221	60-64	Female	2: Medium	Long-term resident	1,155	55.04 (52.63, 57.67)	0.68 (-1.71, 3.06)
2222	60-64	Female	2: Medium	Immigrant	45	65.58 (62.65, 68.50)	-0.04 (-3.28, 3.20)
2231	60-64	Female	3: Low	Long-term resident	1,460	50.13 (47.89, 52.56)	0.38 (-2.00, 2.77)
2232	60-64	Female	3: Low	Immigrant	100	61.67 (58.49, 64.54)	0.25 (-3.07, 3.56)
3111	65-69	Male	1: High	Long-term resident	960	48.01 (45.26, 50.68)	0.03 (-2.59, 2.64)

Table 6.3.3 (continued). Secondary analysis: strata frequencies, predicted probabilities, and intersectional effects
Study period: 2010-2013, N=54,735.

Strata ID	Age	Sex	Income tercile	Immigrant status	N	Predicted 1-year survival probability in % (95% CI) - model 3B	Difference in predicted from expected 1-year survival probability in % (95% CI) - model 3B
3112	65-69	Male	1: High	Immigrant	30	59.64 (56.22, 62.52)	0.00 (-3.69, 3.69)
3121	65-69	Male	2: Medium	Long-term resident	1,725	43.62 (41.33, 45.80)	-0.90 (-3.53, 1.73)
3122	65-69	Male	2: Medium	Immigrant	55	56.02 (52.81, 59.15)	-0.14 (-3.33, 3.04)
3131	65-69	Male	3: Low	Long-term resident	1,950	39.74 (37.48, 41.86)	-0.32 (-2.54, 1.89)
3132	65-69	Male	3: Low	Immigrant	165	51.57 (48.31, 54.68)	-0.22 (-3.28, 2.83)
3211	65-69	Female	1: High	Long-term resident	690	57.16 (54.61, 59.81)	-0.18 (-2.83, 2.46)
3212	65-69	Female	1: High	Immigrant	10	68.31 (65.18, 71.19)	-0.10 (-2.97, 2.78)
3221	65-69	Female	2: Medium	Long-term resident	1,570	55.02 (52.69, 57.30)	1.19 (-1.20, 3.58)
3222	65-69	Female	2: Medium	Immigrant	25	65.19 (62.22, 67.91)	0.00 (-2.76, 2.76)
3231	65-69	Female	3: Low	Long-term resident	2,025	49.35 (47.08, 51.42)	-0.10 (-2.32, 2.13)
3232	65-69	Female	3: Low	Immigrant	125	61.44 (58.47, 64.28)	0.26 (-2.66, 3.18)
4111	70-74	Male	1: High	Long-term resident	755	47.53 (44.92, 50.29)	0.83 (-1.97, 3.63)
4112	70-74	Male	1: High	Immigrant	15	58.41 (55.03, 61.65)	0.00 (-3.24, 3.23)
4121	70-74	Male	2: Medium	Long-term resident	1,860	43.46 (41.34, 45.75)	0.13 (-2.14, 2.40)
4122	70-74	Male	2: Medium	Immigrant	35	54.85 (51.69, 58.10)	-0.23 (-3.54, 3.09)
4131	70-74	Male	3: Low	Long-term resident	2,250	39.45 (37.48, 41.66)	0.66 (-1.75, 3.06)
4132	70-74	Male	3: Low	Immigrant	190	50.95 (47.91, 53.95)	0.37 (-3.02, 3.76)
4211	70-74	Female	1: High	Long-term resident	585	55.84 (53.17, 58.58)	-0.47 (-3.00, 2.05)
4221	70-74	Female	2: Medium	Long-term resident	1,535	53.13 (50.88, 55.60)	0.37 (-1.77, 2.51)
4222	70-74	Female	2: Medium	Immigrant	20	64.05 (61.01, 66.99)	-0.23 (-3.36, 2.89)
4231	70-74	Female	3: Low	Long-term resident	2,395	47.43 (45.47, 49.69)	-0.68 (-2.78, 1.42)
4232	70-74	Female	3: Low	Immigrant	140	59.68 (56.51, 62.53)	-0.16 (-3.12, 2.81)
5111	75-79	Male	1: High	Long-term resident	635	42.96 (40.24, 45.46)	-0.05 (-2.89, 2.80)
5121	75-79	Male	2: Medium	Long-term resident	1,700	40.56 (38.22, 43.10)	0.88 (-1.61, 3.36)
5122	75-79	Male	2: Medium	Immigrant	30	50.98 (47.79, 54.08)	-0.46 (-3.62, 2.70)
5131	75-79	Male	3: Low	Long-term resident	2,140	36.33 (34.32, 38.50)	0.80 (-1.47, 3.07)

Table 6.3.3 (continued). Secondary analysis: strata frequencies, predicted probabilities, and intersectional effects
Study period: 2010-2013, N=54,735.

Strata ID	Age	Sex	Income tercile	Immigrant status	N	Predicted 1-year survival probability in % (95% CI) - model 3B	Difference in predicted from expected 1-year survival probability in % (95% CI) - model 3B
5132	75-79	Male	3: Low	Immigrant	210	46.82 (43.91, 50.05)	-0.02 (-3.44, 3.41)
5211	75-79	Female	1: High	Long-term resident	490	52.42 (49.53, 55.00)	-0.13 (-3.00, 2.74)
5221	75-79	Female	2: Medium	Long-term resident	1,280	48.33 (45.76, 50.72)	-0.87 (-3.43, 1.70)
5222	75-79	Female	2: Medium	Immigrant	20	60.76 (57.57, 63.80)	-0.04 (-3.33, 3.25)
5231	75-79	Female	3: Low	Long-term resident	2,215	44.95 (42.80, 47.10)	0.26 (-2.25, 2.76)
5232	75-79	Female	3: Low	Immigrant	145	56.27 (52.98, 59.43)	-0.15 (-3.27, 2.97)
6111	≥ 80	Male	1: High	Long-term resident	670	35.19 (32.58, 37.93)	-0.13 (-2.74, 2.47)
6112	≥ 80	Male	1: High	Immigrant	15	46.77 (43.05, 50.01)	-0.08 (-3.72, 3.56)
6121	≥ 80	Male	2: Medium	Long-term resident	1,765	32.63 (30.63, 34.89)	0.42 (-1.72, 2.57)
6122	≥ 80	Male	2: Medium	Immigrant	45	43.08 (39.93, 46.10)	-0.26 (-3.65, 3.13)
6131	≥ 80	Male	3: Low	Long-term resident	2,365	30.05 (28.15, 31.90)	1.59 (-0.41, 3.59)
6132	≥ 80	Male	3: Low	Immigrant	280	38.22 (35.56, 41.27)	-0.60 (-3.63, 2.44)
6211	≥ 80	Female	1: High	Long-term resident	500	44.27 (41.58, 47.03)	-0.34 (-3.06, 2.39)
6221	≥ 80	Female	2: Medium	Long-term resident	1,305	40.82 (38.25, 43.22)	-0.24 (-2.89, 2.40)
6222	≥ 80	Female	2: Medium	Immigrant	10	52.73 (49.54, 56.11)	-0.12 (-3.36, 3.11)
6231	≥ 80	Female	3: Low	Long-term resident	2,555	36.92 (34.83, 39.09)	0.06 (-1.99, 2.10)
6232	≥ 80	Female	3: Low	Immigrant	210	47.46 (44.37, 50.37)	-0.98 (-3.96, 2.01)
4212	70-74	Female	1: High	Immigrant	35 [A]	67.39 (64.21, 70.33)	0.02 (-3.09, 3.13)
5112	75-79	Male	1: High	Immigrant		54.75 (51.37, 58.15)	-0.14 (-3.58, 3.29)
5212	75-79	Female	1: High	Immigrant		63.84 (60.58, 66.98)	-0.20 (-3.52, 3.11)
6212	≥ 80	Female	1: High	Immigrant		56.39 (52.96, 59.94)	0.21 (-3.38, 3.79)

[A] Frequencies were collapsed for these strata due to cell counts of less than 10 for each, which cannot be reported as per Statistics Canada's data confidentiality guidelines.

References for appendices

(1) The International Agency for Research on Cancer (IARC). International rules for multiple primary cancers (ICD-0 third edition). 2004 Aug. Report No.: 0959-8278 (Print) 0959-8278 Contract No.: 4.

(2) Johnson CH, Peace S, Adamo P, Fritz A, Percy-Laurry A, Edwards BK. The 2007 multiple primary and histology coding rules. Bethesda, MD: National Cancer Institute, Surveillance, Epidemiology and End Results Program; 2007.

(3) Fritz AG. International classification of diseases for oncology : ICD-O. Third ed. Geneva: World Health Organization; 2000.

(4) Slee VN. The international classification of diseases: ninth revision (ICD-9). American College of Physicians; 1978. p. 424-6.

(5) Percy C, Holten Vv, Muir CS, Organization WH. International classification of diseases for oncology. Second ed. Geneva: World Health Organization; 1990.

(6) Canadian Cancer Society's Advisory Committee. Canadian Cancer Statistics: A 2020 special report on lung cancer. Toronto, ON: Canadian Cancer Society; 2020.

(7) Pinard D. Methodology Changes: Census Family Low Income Measure Based on the T1 Family File: Statistics Canada; 2018.

(8) Qureshi S, Schnitzer ME, Nandi A, Strumpf E. Social inequalities in survival among Canadian lung cancer patients and the mediating role of stage at diagnosis. 2024.

(9) Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43(11):1130-9.

(10) Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40(5):373-83.

(11) Gasparini A. comorbidity: An R package for computing comorbidity scores. Journal of Open Source Software. 2018;3(23):648.

(12) Postal Code OM Conversion File Plus (PCCF+) [Internet]. Government of Canada. 2023 [cited 2023-12-06]. Available from: https://www150.statcan.gc.ca/n1/en/catalogue/82F0086X#wb-auto-2.

(13) Wilkins R. PCCF+ Version 5E User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through March 2009. In: Health Analysis Division, editor. Ottawa, ON.2009.

Chapter 7. Discussion and Conclusion

7.1. Summary of findings

The third SDG seeks to ensure healthy lives for all by 2030, aiming to reduce premature mortality from non-communicable diseases, including lung cancer, by one-third through prevention, treatment, and universal access to healthcare and treatments.[1] This goal aligns with the CPAC 2019–2029 strategic plan to reduce Canada's cancer burden by improving access to timely diagnoses and high-quality treatments while eliminating inequities in cancer care.[10] Achieving these objectives requires a commitment to proportional universalism—universal interventions tailored to diverse needs—and, fundamentally, the identification of populations facing poorer cancer outcomes and barriers to equitable care. My thesis supports these goals by identifying key populations in Canada that may be experiencing inequities in lung cancer survival and care.

In **Manuscript 1**, I documented inequalities in lung cancer across measures of individual-level SES (income and education) and by place of residence among respondents of the CCHS surveys who were diagnosed with lung cancer between 2010 and 2016 in Canada. As expected, individuals from the lowest SES groups had lower hazards of death compared to individuals from the highest SES groups. Interestingly, when analyses were stratified by sex, income-based inequalities in survival were more prominent in women. Similarly, immigrant women experienced worse survival than non-immigrant women, although no survival differences were observed by immigrant status in the overall study sample. Education-based inequalities in lung cancer survival did not vary much across males and females, indicating that education may be capturing elements of SES that are different from income.

The largest differences in the hazard of death were observed between individuals from large and small urban cities, with an estimated difference of 50.1 deaths per 10,000 person-months (95% CI: -3.7 to 107.2). Mediation analyses revealed that approximately 41% of this effect was mediated by stage at diagnosis (stages IIIb and IV). In a post-hoc analysis, in which individuals diagnosed prior to responding to the CCHS were excluded and advanced stage was redefined by combining stage IIIa with stages IIIb and IV, the hazard difference between small and large urban cities did not change: 48.5 deaths per 10,000 person-months (95% CI: 12.1 to 87.5). However, the proportion

of this difference found to be mediated by advanced stage at diagnosis increased substantially to 76%.

Without widespread lung cancer screening in Canada, most patients present with severe symptoms or symptoms suspicious of lung cancer, which are typically associated with advanced-stage diagnoses, emergency care, and expedited investigations.[238-241] In contrast, patients with less severe or specific symptoms are more likely to have early-stage cancer (stages I and II) and seek care outside emergency settings, both of which are associated with longer diagnostic delays.[93, 240] These delays may be further compounded by health system factors, such as resource shortages, resulting in individuals with early-stage lung cancer having their disease progress to later stages by the time the investigation is completed.

The larger difference in stage IIIa lung cancers observed between small and large urban cities (17% vs. 13%) strengthens the plausibility that survival inequalities between patients from these areas are driven by health system factors that cause long investigation delays and advanced stages at diagnosis. Research shows that individuals in small urban and rural areas in Canada are less likely to consult specialists.[242, 243] In Ontario, individuals from urban health regions outside of Toronto, which include small urban cities, are twice as likely to face challenges accessing specialists compared to Toronto residents.[244] A longer wait time is the main barrier to seeing a specialist for individuals in urban health regions outside of Toronto. It is, therefore, possible that wait times for specialists increase with decreasing urbanicity, potentially contributing to longer diagnostic delays for lung cancer and a higher proportion of stage IIIa diagnoses in small compared to large urban areas in Canada.

Although methods exist to account for overall survival when studying inequalities, researchers often report survival outcomes at specific time points during follow-up, such as median survival time or 1-, 3-, or 5-year survival probabilities. However, this approach has its challenges. For instance, lung cancer has significantly lower survival rates compared to other cancers, largely because it predominantly affects seniors and is most often diagnosed at advanced stages. The 5-year lung cancer survival rate is only 22%,[15] so comparing survival between groups at 5 years focuses on the fittest individuals who survive the longest. As a result, long-term survivors may differ significantly from the rest of the population, and inequalities between groups may appear

reduced because the remaining survivors become more similar over time. Similarly, when survival probabilities are compared at 1 year—which corresponded to the median survival time in Manuscript 1—differences between groups may appear minimal. This is because more than half of lung cancer patients are diagnosed at advanced stages when the disease is so lethal that it may overshadow potential inequalities between groups due to other factors.

Analyzing survival across the full follow-up period can uncover inequalities that may otherwise go unnoticed and offer insights into factors that potentially mediate these disparities. This is exemplified in Manuscript 1, where the median survival time between small and large urban centers showed minimal differences. This occurred because, in both groups, more than 50% of individuals were diagnosed at advanced stages. Since advanced-stage diagnosis was found to be the major factor driving inequalities between these groups, it makes sense that we did not observe differences in survival during the early phase of follow-up, such as at the median survival time. These observations further reinforce our conclusion regarding the important mediating role of stage at diagnosis in explaining survival inequalities between large and small urban centers. Moreover, despite more than 50% of individuals in each socioeconomic group being diagnosed at advanced stages, the observed differences in median survival time ranged from 3 to 4 months between the highest and lowest levels of SES. This suggests that the effect of advanced stage at diagnosis on survival may vary by SES or that additional factors mediate survival inequalities beyond advanced stage at diagnosis.

Manuscript 1 also highlights the importance of examining absolute differences in hazards of death, rather than relying solely on hazard ratios, when assessing survival across groups. Given lung cancer's high case-fatality rate (88 deaths per 100 people annually),[245] baseline death rates are expected to be high, making meaningful hazard differences appear small on the ratio scale. For example, hazard rates of 775 and 725 deaths per 10,000 person-months yield a difference of 50 deaths per 10,000 person-months, but a hazard ratio of just 1.1. This may explain why many studies comparing urban and rural areas report hazard ratios between 0.9 and 1.1.[93, 96, 98, 246, 247] Dichotomous urban-rural measures used in some studies may have further diluted meaningful differences.

Although confidence intervals were wide, results from Manuscript 1 suggest that individuals in rural areas with weak or no metropolitan influence have better survival than those in large urban areas, aligning with Ontario-based studies that reported hazard ratios between 0.97 and 1.[93, 98, 247] A Quebec study also found a hazard ratio of 0.97 when comparing remote rural to urban regions.[246] While these estimates were not statistically significant, they align with previous studies demonstrating that rural lung cancer patients are less likely to die within 30 days of diagnosis,[94] are less likely to be diagnosed at advanced stages,[248] and are more likely to receive timely surgery.[249] In remote rural areas of Ontario, one study showed that patients were more likely to access both surgery and adjuvant chemotherapy compared to urban areas.[28] In contrast, two nationwide studies reported slightly lower rates of curative surgery among rural versus urban dwellers,[97, 250] and two studies reported lower survival rates in rural areas relative to urban areas in Alberta.[96, 251] Overall, more research is needed to understand how geography-based survival inequalities vary across provinces, time, and by stage at diagnosis.

In **Manuscript 2,** although individual-level measures of SES were not available for analysis, differences in lung cancer survival were found by area-based measures of SES among individuals receiving targeted therapies in Quebec between 2001 and 2019. These findings underscore the importance of therapeutic advancements for advanced lung cancer, which are contributing to its reduced lethality but also to the emergence of inequalities in survival. Given that the majority of patients are diagnosed at an advanced stage, this has broader implications for survival inequalities in the overall lung cancer population.

The median survival time ratios were most pronounced when comparing individuals from areas with the lowest and highest material deprivation index, though the relative difference in receiving targeted therapies as second-line treatment was not the largest across these groups. One potential explanation for this is that 53% of individuals from the least deprived areas also resided in Montreal, where the likelihood of receiving a targeted therapy was 2.6 (95% CI: 1.4-6.3) times that of individuals living in other large urban areas. This also suggests that access to second-line targeted therapies may not fully explain SES-based survival differences. The fact that Montreal patients relative to other areas were more likely to receive therapies as second-line treatment aligns with patterns of implementation of genetic testing for targeted therapies, as the first two centers in Quebec to offer these tests were located in Montreal.[252] Moreover, one of these centers began

reflex testing in 2017, whereby genetic testing of tumours was performed at the same time as tumours were assessed to confirm cancer diagnosis.[253] This suggests that access to targeted therapies may have improved over time not only due to their inclusion in public drug formularies but also because of the increased availability of genetic testing. However, secondary analyses in Manuscript 2 revealed that access to targeted therapies over time remained higher in Montreal compared to other urban or rural areas but also became much lower in health centers in intermediary, remote, or northern regions relative to health regions with university-affiliated hospitals.

Unlike in the previous manuscript, lung cancer patients from the most immigrant-dense areas had better survival compared to those from the least immigrant-dense areas. This finding could reflect the "healthy immigrant effect," as more densely populated immigrant areas may be representative of recent immigrants, who tend to exhibit better health outcomes upon arrival. Immigrants often initially settle in immigrant-dense, ethnically diverse neighbourhoods but tend to relocate to more affluent, predominantly White areas as they accumulate wealth and become long-term residents.[254] Another potential explanation for this survival advantage is that 90% of individuals living in the most immigrant-dense areas were from Montreal, where the likelihood of receiving targeted therapies was relatively high compared to other urban or rural areas.

Finally, delays in receiving targeted treatments as first-line palliative treatment were shortest in Quebec's health and social services regions that included university-affiliated hospitals (Quebec, Montreal, and Sherbrooke). However, individuals from Montreal, compared to those from other urban areas, had a significantly longer median wait time for receiving targeted therapies. This indicates that significantly shorter wait times for individuals from health regions with university-affiliated hospitals may be largely driven by shorter wait times at the Institut universitaire de cardiologie et de pneumologie de Québec – Université Laval (IUCPQ-UL) in Quebec. IUCPQ-UL has cut median wait times from referral to treatment by approximately half relative to the rest of Quebec (26 vs. 56 days) due to standardized protocols in their pathology lab for diagnosis and genetic testing of tumour samples.[255] The consistent differences in wait times between health regions with university-affiliated hospitals and those peripheral to Montreal or Quebec urban cores throughout different population quantiles indicate that delays are most likely due to health system-level barriers and are inequitable. Although survival differences were not observed between these

health regions, the longer delays in receiving treatment in regions peripheral to Quebec or Montreal urban cores may compound delays already experienced by certain individuals. For example, older patients and immigrants have been found to face longer times between their first healthcare encounter and lung cancer diagnosis.[93]

In secondary analyses of the first manuscript, the effects of low income seemed stronger among women than men. In the second manuscript, survival differences were most pronounced using the material deprivation index—a composite measure of low income, unemployment, and low education. These findings highlight how social factors could interact and shape lung cancer survival, providing the rationale for **Manuscript 3**. In this manuscript, I focused on an intersectional analysis of age, sex, income, and immigrant status to understand how multiple axes of advantage or disadvantage influence lung cancer survival in Canada (2002–2013). Results revealed large variations in 1-year survival across intersectional strata (23% to 67%). Notably, low-income men aged <60 who were long-term residents had one of the lowest survival probabilities (37%), comparable to low-income men aged 70–74 (34%) and lower than all other strata aged <60. Their survival was also significantly lower than expected based on the additive effects of age, sex, income, and immigrant status. In contrast, low-income women aged ≥80 who were immigrants generally had the highest survival probabilities among all strata with the same age group. However, they had lower-than-expected 1-year survival based on additive effects alone. This resulted in low-income immigrant women experiencing the largest gap in survival when comparing the oldest to the youngest strata, relative to the gap due to age among low-income long-term resident women, low-income immigrant men, and low-income long-term resident men. Overall, these findings exemplify that the effect of age does not operate independently of other social factors, such as low income, sex, and immigrant status, or vice versa, and that the effects of these social identities are mutually constitutive.

Although I discuss smoking as a potential explanation for the lower-than-expected survival among low-income men aged <60 who were long-term residents, the intent of this analysis is not to further marginalize these individuals. Instead, it aims to highlight the broader marginalizing conditions they may face due to smoking, such as stigma associated with smoking, reduced access to cessation programs, and structural inequities in healthcare access and support systems.[225, 256, 257] For low-income immigrant women, the healthy-immigrant effect may diminish with age, contributing

to lower-than-expected survival among the oldest long-term immigrant women.[160] The exact reasons for these intersectional effects were not investigated in this work. However, secondary analyses do suggest that stage at diagnosis likely accounts for many of these deviations, particularly for low-income men who are long-term residents. More data and formal mediation analyses are needed to investigate this hypothesis.

While only a few intersectional effects were observed for 1-year survival, this does not rule out the presence of intersectional effects for other cancer outcomes or at later time points. As discussed earlier, 1-year survival differences may appear minimal between groups where most individuals are diagnosed at advanced stages, potentially obscuring inequalities. These inequalities may become more apparent at longer follow-up times, such as two years, which could also uncover intersectional effects.

7.2. Strengths and limitations

There are several limitations to the work in this thesis that warrant careful consideration, in addition to those outlined in each manuscript. First, the datasets used in this study may not fully represent all individuals in Canada affected by lung cancer. For example, in Manuscript 1, the CCHS does not include individuals residing in institutions such as long-term care. Given that lung cancer predominantly affects seniors, this exclusion likely omits a portion of the lung cancer population. Similarly, hard-to-reach populations, such as individuals experiencing homelessness, may not have been adequately captured in all three studies. Homeless individuals often lack a fixed address, which complicates their inclusion in surveys like the CCHS, which relies on respondents residing in private or collective dwellings. Additionally, homeless individuals are less likely to have a health card or file taxes due to their lack of a permanent address, making them more likely to be excluded from our analyses due to missing postal code in health administrative data and/or missing income data when linked to tax records. This underrepresentation is important because approximately 80% of homeless individuals in Canada are smokers,[258] a strong risk factor for lung cancer. Homeless individuals are four times more likely to develop cancer than the general population.[259] In the United States, homeless adults aged 45–64 have cancer mortality rates two to three times higher than the general population, with cancer as the leading cause of death and lung cancer as the primary cause of cancer deaths.[260] In Canada, about 30 to 40% of the

homeless population is represented by individuals aged ≥45.[261] As the homeless population in Canada continues to grow overall, and among seniors,[262] greater efforts are needed to include this marginalized group in research and data collection.

Another important group of individuals that need further consideration are the Indigenous populations in Canada, who experience significantly higher rates of smoking, as well as lung cancer incidence, mortality, and survival relative to the non-Indigenous population.[263] Indigenous people are also more likely to experience low-income and homelessness.[97, 261] In general, race/ethnicity has been shown to predict lung cancer survival in the United States.[264] Although the CCHS collects information on race/ethnicity, the majority (approximately 90%) of individuals in the study for Manuscript 1 identified as White. As a result, the number of respondents from other racial and ethnic groups was too small to estimate survival differences by race/ethnicity. Given the diversity of Canada's immigrant population, it is important to analyze immigration status in conjunction with race/ethnicity.

Differences observed between Manuscripts 1, 2, and 3 regarding survival disadvantages or advantages among immigrants may partly reflect variations in time since immigration. Manuscript 1 likely includes immigrants who have lived in Canada for a longer period, whereas Manuscripts 2 and 3 may focus on more recent immigrants. Additionally, these differences could be influenced by temporal shifts in the regions from which immigrants are admitted. These findings highlight the importance of examining intersectional effects involving race/ethnicity, immigrant status, and sex/gender.

The linkage rates of the datasets used in Manuscripts 1 and 3 to the DRD at Statistics Canada were all above 95%, indicating that most individuals from each dataset were successfully linked. Furthermore, between-dataset linkage rates—for example, linking CCHS or CCR to the T1FF—were as high as 98%. Linkage rates with other administrative databases, including DAD, NACRS, CVSD, T1PMF-D, and IMDB, were generally lower since not all individuals are in the study cohorts are expected to be hospitalized or seek medical care, experience death, or be immigrants. In Manuscript 1, only about 10% of CCHS respondents could be linked to the CCR or CVSD. In Manuscript 3, the linkage rates of CCR to data from DAD and NACRS from 2005–2008, the period used for calculating comorbidities, were approximately 70% and 50%, respectively. The

completeness of these datasets has improved over time, with linkage rates rising to about 90% for DAD and 80% for NACRS in the period 2011–2015, which fall outside our period of interest. This suggests that 20–30% of individuals classified as having no comorbidities in our study may, in fact, have had uncaptured conditions, limiting our ability to fully assess the contribution of comorbidities to observed intersectional effects. The linkage of CCR to IMDB was approximately 6% overall but increased from 2% to 11% between 1992 and 2013. This reaffirms that not all immigrants were captured through IMDB and that the likelihood of identifying immigrants with lung cancer—particularly those who arrived more than 30 years ago—increases with longer study periods.

Linkage rates for the lung cancer cohort in Manuscript 2 with other datasets were not available. However, the lung cancer cohort is expected to be population-representative, as it includes all individuals with public health insurance coverage, which is universal in Quebec. In contrast, the drug prescription database (SMED) includes only individuals covered by the public prescription drug insurance plans, which are not universal. As a result, linkage between the lung cancer cohort and SMED likely excludes individuals with private drug insurance. As discussed in the manuscript, these individuals are more likely to be younger and healthier, and their exclusion may bias estimates of socioeconomic and sociodemographic inequalities in lung cancer survival toward the null. However, this bias is estimated to be low since lung cancer mainly affects seniors (≥ 65 years), who in turn are mainly (90%) are covered by a public drug insurance in Quebec.[216, 265]

Finally, there are some limitations regarding the exposures and outcomes used in my analyses that need to be addressed. In Manuscript 2, I was restricted to using area-level measures of income, education, and immigration status, as the linked Quebec health administrative databases used for the study were situated outside the RDC and not linked to Statistics Canada datasets containing individual-level information. Area-level measures of income have been shown to be imperfect proxies for individual-level income, often underestimating survival differences based on individual income.[266] This may also apply to area-level measures of education and immigration status. However, identifying disparities in cancer outcomes at the area level remains valuable, as it can highlight inequities that reflect both individual and neighbourhood-level disadvantages, which could potentially be addressed simultaneously through neighbourhood-level policies and interventions. Furthermore, I refer to household income as a measure of individual-level income

in this thesis, with the assumption that resources are shared within a household. However, income-sharing can be complex, where it is more applicable for couples, but can be less applicable to seniors living with other family members who may be benefitting from their income and depleting them of resources.[267] This could lead to an underestimation of survival differences if, for instance, individuals from low-income households experience greater financial autonomy by living alone, while those from high-income households may have less autonomy due to residing in intergenerational homes.

In terms of outcome, in all three studies, I opted to assess deaths from all causes rather than lung cancer-specific deaths. The limitation of this approach is that when differences in survival are observed between two groups, it is difficult to determine if longer survival in one group is caused by fewer deaths from lung cancer or from any cause. Assessing lung cancer-specific survival may be more relevant. However, this approach censors individuals who die from other causes, implicitly assuming that lung cancer is the only possible cause of death. This assumption may not be appropriate for lung cancer, a disease that primarily affects seniors and individuals with a smoking history, both of whom are at high risk of dying from other causes. Furthermore, cancer-specific survival is known to overestimate survival for cancers with poor prognoses or those frequently diagnosed at advanced stages, such as lung cancer, due to the miscoding of cancer-related deaths as deaths from other causes.[268] Due to its poor prognosis, previous studies on lung cancer have generally reported minimal differences between relative estimates based on all-cause survival and those based on cancer-specific survival.[269-271]

This thesis demonstrates several strengths in both data and methodology. In Manuscripts 1 and 3, I used linked datasets with individual-level measures of income, education, and immigration status to examine lung cancer survival, providing a national overview of inequalities among underrepresented populations in research. These datasets enabled the application of advanced analytical techniques, such as intersectional and mediation analyses, to uncover nuanced survival patterns and identify key drivers of inequalities, such as stage at diagnosis. Although Quebec was excluded by design due to my focus on advanced stage at diagnosis, I addressed this gap in Manuscript 2 by utilizing more recent health administrative databases from Quebec. The linked datasets used in Manuscript 2 were distinct from the linked datasets I used in the other studies as they included prescription claims data. This enabled me to examine socioeconomic,

sociodemographic, and geographic inequalities in survival, revealing persistent disparities even among advanced cancer patients using take-home cancer drugs—a treatment modality that is becoming increasingly common.

A unique strength of the income measure in Manuscripts 1 and 3 is the use of the full cohort of CCHS respondents, linked to tax files. This enabled the creation of national-level income tertiles for each year, adjusted by province and urban-rural area to account for temporal trends and cost-of-living differences. These refined income tertiles provided a consistent, population-level reference for the individuals in my studies, allowing for a more accurate assessment of income-based disparities in lung cancer survival across Canada. By assessing place of residence in several ways in the three manuscripts, I was able to uncover important differences within urban areas with respect to cancer survival. Immigrant status was examined across all three manuscripts, providing valuable insights despite differing results. These variations allowed me to speculate that survival differences among immigrants may depend on factors such as time since immigration and ethnicity, highlighting the need for further exploration of these intersecting influences.

A unique strength of the analyses in this thesis is that they reflect the complexities inherent in survival studies, particularly in Manuscript 1. I noted differences in results depending on whether I focused on median survival time or hazard differences and highlighted the importance of time dependence in survival analysis. Specifically, the indirect effects during the mediation analysis were influenced by time-dependent hazard differences, providing deeper insight into the mediating role of stage at diagnosis. Finally, while not directly demonstrated, I highlight in section 7.1 of this thesis the implications of using absolute versus relative survival measures, emphasizing how different approaches can lead to distinct conclusions about survival outcomes.

7.3. Implications

In the United States and other high-income European countries, cancer registries are supplemented with relevant data, such as race/ethnicity, place of birth, and education, to study inequities in cancer outcomes at the national level.[272-275] Furthermore, nationally representative surveys have been linked to cancer registries.[274-276] Hence, different data sources have enabled the documentation of population-level social inequalities in outcomes across various stages of the cancer care continuum—prevention, diagnosis, treatment, and survival—in these populations.[277, 278] In

contrast, studying inequalities in cancer outcomes is more challenging in Canada, as the CCR does not incorporate information on SES, race/ethnicity, or place of birth. While other databases containing such information exist, their routine linkage to the CCR is needed.[279] This limits the ability to comprehensively assess social inequalities in cancer outcomes, which is a fundamental step in identifying potential inequities and addressing them effectively.

In this thesis, I highlight the potential of non-traditional data sources, when linked to health administrative databases, to advance research on cancer disparities. The CCHS, T1FF, and IMDB are rich datasets that remain underutilized for this purpose. However, I encountered challenges, including outdated data and small sample sizes in certain strata, such as immigrants. For instance, in Manuscript 3, I could only include lung cancer cases diagnosed up to 2013, leaving a decade-long gap. Routine updates to these linkages are essential to ensure that health equity research remains both relevant and impactful. Furthermore, while these datasets are accessible through Statistics Canada, they lack detailed information on prescription drug utilization. In my second manuscript, I demonstrate the value of incorporating prescription claims data to study inequalities in cancer outcomes. Linking the CCR with drug data (e.g., public and private drug insurance claims data, community and hospital pharmacy data) and other complementary datasets could enable national-level inferences based on individual-level data. Such integration would provide valuable insights as at-home cancer treatments become increasingly common but remain unequally accessible across populations.

One of the most important findings from my research is the identification of significant differences in survival within urban areas, demonstrating that not all urban settings are the same. Despite large urban areas being more populous, which allows for more granular analyses, researchers often continue to treat them as a single, homogeneous group. This finding underscores the need for more nuanced approaches to urban research. However, this should not detract from the importance of studying rural areas, which are known to face unique challenges and disparities. Furthermore, I have identified key populations who may be facing unique challenges in surviving lung cancer, such as young long-term resident men and immigrant women, specifically those who have been in Canada for an extended period. I have also found important differences in stage at diagnosis and treatment, in particular by place of residence.

The data used in this thesis were collected before the COVID-19 pandemic, which disrupted Canada's healthcare system, created a backlog, and likely exacerbated existing health inequalities, including those related to cancer outcomes.[184, 280-283] As a result, the populations identified in this thesis may have been disproportionately affected, further widening lung cancer survival disparities. However, many of these disparities could be mitigated through early diagnosis and improved access to treatments. More research is needed to better understand these populations and identify effective interventions. This is especially critical in the evolving landscape of lung cancer care in Canada, where resource constraints persist. Among all Organization for Economic Cooperation and Development (OECD) countries, Canada has the lowest number of diagnostic imaging equipment, beds, and oncologists per population. As newer, costly therapies are added to public drug coverage plans and provinces implement lung cancer screening programs with targeted outreach, applying the principle of universal proportionalism—allocating resources to maximize benefits for those most in need—will be key to reducing, rather than deepening, disparities in lung cancer outcomes.

7.4. Conclusion

Although lung cancer is the leading cause of cancer-related deaths in Canada, it remains critically underfunded in research, largely due to the stigma associated with smoking and the lack of long-term survivors advocating for increased support.[284, 285] However, investing in data-linkage efforts for continuous and timely analyses is critical for identifying and monitoring at-risk groups, developing targeted interventions, and achieving global and local targets for reducing inequities in lung cancer care and mortality.

In this thesis, I have demonstrated the importance of leveraging linked data to understand survival disparities across populations with different social positions. By incorporating downstream social determinants of health, such as diagnosis and treatment, my analyses provide valuable insight into potential avenues for mitigating survival inequalities. Through an intersectional approach, I also illustrate how multiple axes of social identity—such as age, sex, income, and immigration status—can intertwine to shape survival experiences. This highlights the need for future research to move beyond examining each social identity in isolation. Most importantly, I have identified social inequalities in survival overall and at advanced stages of lung cancer. Key populations found to

experience survival disadvantages but not adequately addressed in previous lung cancer research include individuals living in small urban cities, low-income young men who are long-term residents, and immigrant women who have been in Canada for many decades. These populations require special attention in future research efforts to reduce lung cancer survival disparities and, in turn, improve diagnosis, treatment, and overall survival outcomes.

References

1. United Nations. Resolution adopted by the General Assembly on 25 September 2015 (A/RES/70/1): Transforming Our World: The 2030 Agenda for Sustainable Development. New York, United States: United Nations; 2015.

2. Whitehead M. The concepts and principles of equity and health. *Int J Health Serv*. 1992;22(3):429-45. DOI:

3. Public Health Agency of Canada. Key health inequalities in Canada: a national portrait. Government of Canada; 2018.

4. Bray F, Laversanne M, Weiderpass E, Soerjomataram I. The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer*. 2021;127(16):3029-30. DOI: https://doi.org/10.1002/cncr.33587

5. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229-63. DOI: https://doi.org/10.3322/caac.21834

6. Mahase E. Cancer overtakes CVD to become leading cause of death in high income countries. *BMJ*. 2019;366:l5368. DOI: 10.1136/bmj.l5368

7. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71(3):209-49. DOI: https://doi.org/10.3322/caac.21660

8. Government of Canada. Taking Action Together: Canada's 2021 Annual Report on the 2030 Agenda and the Sustainable Development Goals. In: Sustainable Development Goals Unit, editor. Quebec, Canada: Government of Canada,; 2022.

9. United Nations Global indicator framework for the sustainable development goals and targets of the 2030 agenda for sustainable development. In: Division S, editor. 2019. p. 24.

10. Canadian Partnership Against Cancer. 2019-2029: Canadian Strategy for Cancer Control. Doing together what cannot be done alone.; 2019.

11. Dunn J. It Is Time to Close the Gap in Cancer Care. *JCO Glob Oncol*. 2023;9:e2200429. DOI: 10.1200/go.22.00429

12. Union for International Cancer Control. Close the care Gap: World Cancer Day 2022-2024 Campaign Summary Report.; 2024.

13. Schroeder SA, Koh HK. Tobacco Control 50 Years After the 1964 Surgeon General's Report. *JAMA*. 2014;311(2):141-3. DOI: 10.1001/jama.2013.285243

14. Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. *Eur Respir J*. 2016;48(3):889-902. DOI: 10.1183/13993003.00359-2016

15. Canadian Cancer Society's Advisory Committee in collaboration with the Canadian Cancer Society, Statistics Canada and the Public Health Agency of Canada. Canadian Cancer Statistics 2023. Toronto, ON: Canadian Cancer Society; 2023.

16. Corsi DJ, Boyle MH, Lear SA, Chow CK, Teo KK, Subramanian SV. Trends in smoking in Canada from 1950 to 2011: progression of the tobacco epidemic according to socioeconomic status and geography. *Cancer Causes Control*. 2014;25(1):45-57. DOI: 10.1007/s10552-013-0307-9

17. Cancer Care Ontario. Cancer Fact: Lung cancer mortality differences between men and women influenced by smoking trends. 2015.

18. Khuder SA. Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. *Lung Cancer*. 2001;31(2-3):139-48. DOI: 10.1016/s0169-5002(00)00181-1

19. 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;17(6):1471-4. DOI: https://doi.org/10.1245/s10434-010-0985-4

20. Thomas KW, Gould MK. Tumor, Node, Metastasis (TNM) staging system for lung cancer 2020 [Available from: https://www.uptodate.com/contents/tumor-node-metastasis-tnm-staging-system-for-lung-cancer.

21. Canadian Cancer Society's Advisory Committee. Canadian Cancer Statistics: A 2020 special report on lung cancer. Toronto, ON: Canadian Cancer Society; 2020.

22. Darling GE, Tammemägi MC, Schmidt H, et al. Organized Lung Cancer Screening Pilot: Informing a Province-Wide Program in Ontario, Canada. *The Annals of Thoracic Surgery*. 2021;111(6):1805-11. DOI: https://doi.org/10.1016/j.athoracsur.2020.07.051

23. Hilzenrat RA, Yip S, Melosky B, et al. Disparate Time-to-Treatment and Varied Incidence of Actionable Non-Small Cell Lung Cancer Molecular Alterations in British Columbia: A Historical Cohort Study. *Current Oncology*. 2023;30(1):145-56. DOI: https://www.mdpi.com/1718-7729/30/1/12

24. Rami-Porta R, Wittekind C, Goldstraw P. Complete resection in lung cancer surgery: proposed definition. *Lung Cancer*. 2005;49(1):25-33. DOI: https://doi.org/10.1016/j.lungcan.2005.01.001

25. Xie L, Ugnat A-M, Morriss J, Semenciw R, Mao Y. Histology-related variation in the treatment and survival of patients with lung carcinoma in Canada. *Lung Cancer*. 2003;42(2):127-39. DOI: https://doi.org/10.1016/S0169-5002(03)00283-6

26. Woodard GA, Jones KD, Jablons DM. Lung Cancer Staging and Prognosis. In: Reckamp KL, editor. *Lung Cancer: Treatment and Research*. Cham: Springer International Publishing; 2016. p. 47-75.

27. Raptis CA, Bhalla S. The 7th Edition of the TNM Staging System for Lung Cancer: What the Radiologist Needs to Know. *Radiol Clin North Am*. 2012;50(5):915-33. DOI: https://doi.org/10.1016/j.rcl.2012.06.001

28. Evans WK, Stiff J, Woltman KJ, et al. How equitable is access to treatment for lung cancer patients? A population-based review of treatment practices in Ontario. *Lung cancer management*. 2017;6(3):77-86. DOI: 10.2217/lmt-2017-0013

29. Gergen AK, Scott CD, Mitchell JD. Surgery for limited stage small cell lung cancer. *J Thorac Dis*. 2020;12(10):6291-7. DOI: 10.21037/jtd.2020.03.79

30. Dawe DE, Rittberg R, Syed I, et al. Real-world predictors of survival in patients with limited-stage small-cell lung cancer in Manitoba, Canada. *Front Oncol*. 2023;13. DOI: 10.3389/fonc.2023.1191920

31. Shepherd FA, Crowley J, Van Houtte P, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: Proposals Regarding the Clinical Staging of Small Cell Lung Cancer in the Forthcoming (Seventh) Edition of the Tumor, Node, Metastasis Classification for Lung Cancer. *J Thorac Oncol*. 2007;2(12):1067-77. DOI: https://doi.org/10.1097/JTO.0b013e31815bdc0d

32. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. *BMJ*. 1995;311(7010):899. DOI: 10.1136/bmj.311.7010.899

33. Yuan M, Huang LL, Chen JH, Wu J, Xu Q. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduct Target Ther*. 2019;4:61. DOI: 10.1038/s41392-019-0099-9

34. Shokoohi A, Al-Hashami Z, Moore S, et al. Effect of targeted therapy and immunotherapy on advanced nonsmall-cell lung cancer outcomes in the real world. *Cancer Medicine*. 2022;11(1):86-93. DOI: https://doi.org/10.1002/cam4.4427

35. Akhtar-Danesh N, Finley C. Temporal trends in the incidence and relative survival of non-small cell lung cancer in Canada: A population-based study. *Lung Cancer*. 2015;90(1):8-14. DOI: https://doi.org/10.1016/j.lungcan.2015.07.004

36. Ellison LF, Saint-Jacques N. Five-year cancer survival by stage at diagnosis in Canada. *Health Rep*. 2023;34(1):3-15. DOI: https://doi.org/10.25318/82-003-x202300100001-eng

37. Phillips WJ, Leighl NB, Blais N, Wheatley-Price P. Oral targeted therapy for the treatment of non–small cell lung carcinoma. *Can Med Assoc J*. 2024;196(16):E558-E61. DOI: 10.1503/cmaj.231562

38. Qureshi S, Boily G, Boulanger J, et al. Advanced Lung Cancer Patients' Use of EGFR Tyrosine Kinase Inhibitors and Overall Survival: Real-World Evidence from Quebec, Canada. *Curr Oncol*. 2022;29(11):8043-73. DOI: 10.3390/curroncol29110636

39. Ministère de la Santé et des Services sociaux (MSSS). Les partenariats, gege de qualité en cancérologie: plan d'action 2024-2026 du Programme québécois de cancérologie. Gouvernement du Québec; 2023. Contract No.: 978-2-550-97967-8

40. Incidence for Selected Cancers, by Geographic Region, and Sex [Internet]. Canadian Cancer Statistics Dashboard,. 2024. Available from: https://cancerstats.ca/Incidence/Region?la=en#Incidence_for_Selected_Cancers_by_Geographic_Region_and_Sex.

41. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. WHO Document Production Services; 2010.

42. Hahn RA. What is a Social Determinant of Health? Back to Basics. *Journal of Public Health Research*. 2021;10(4). DOI: 10.4081/jphr.2021.2324

43. World Health Organization. *Closing the gap in a generation - Health equity through action on*. Geneva, Switzerland: Atlantic Collaborative on Injury Prevention; 2008. Available from: https://canadacommons.ca/kbart/artifacts/1886481/closing-the-gap-in-a-generation/2635796.

44. Wakefield MK, Williams DR, Le Menestrel S, Flaubert JL, National Academies of Sciences E, Medicine. *The future of nursing 2020-2030 : charting a path to achieve health equity*. Washington, DC: National Academies Press; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK573914/.

45. Glymour MM, Avendano M, Kawachi I. 17Socioeconomic Status and Health. In: Berkman LF, Kawachi I, Glymour MM, editors. *Social Epidemiology*: Oxford University Press; 2014. p. 0.

46. World Health Organization. *Handbook on health inequality monitoring: with a special focus on low-and middle-income countries*: World Health Organization; 2013.

47. World Health Organization. *Operational framework for monitoring social determinants of health equity*: World Health Organization; 2024.

48. Pan-Canadian Public Health Network Partners in Public Health. Indicators of Health Inequalities. A report from the Population Health Promotion Expert Group and the Healthy Living Issue Group for the Pan-Canadian Public Health Network.; 2010.

49. Krieger N. Defining and investigating social disparities in cancer: critical issues. *Cancer Causes Control*. 2005;16(1):5-14. DOI: 10.1007/s10552-004-1251-5

50. Quaglia A, Lillini R, Mamo C, Ivaldi E, Vercelli M. Socio-economic inequalities: A review of methodological issues and the relationships with cancer survival. *Critical Reviews in Oncology / Hematology*. 2013;85(3):266-77. DOI: 10.1016/j.critrevonc.2012.08.007

51. Dwyer LL, Vadagam P, Vanderpoel J, Cohen C, Lewing B, Tkacz J. Disparities in Lung Cancer: A Targeted Literature Review Examining Lung Cancer Screening, Diagnosis, Treatment, and Survival Outcomes in the United States. *J Racial Ethn Health Disparities*. 2024;11(3):1489-500. DOI: 10.1007/s40615-023-01625-2

52. Rivera MP, Gudina AT, Cartujano-Barrera F, Cupertino P. Disparities Across the Continuum of Lung Cancer Care. *Clin Chest Med*. 2023;44(3):531-42. DOI: https://doi.org/10.1016/j.ccm.2023.03.009

53. Namburi N, Timsina L, Ninad N, Ceppa D, Birdas T. The impact of social determinants of health on management of stage I non-small cell lung cancer. *The American Journal of Surgery*. 2022;223(6):1063-6. DOI: https://doi.org/10.1016/j.amjsurg.2021.10.022

54. Cui W, Finkelstein J. Using EHR Data to Identify Social Determinants of Health Affecting Disparities in Cancer Survival. *Stud Health Technol Inform*. 2022;290:967-71. DOI: 10.3233/SHTI220224

55. Coaston TN, Sakowitz S, Chervu NL, et al. Social determinants as predictors of resection and long-term mortality in Black patients with non-small cell lung cancer. *Surgery*. 2024;175(2):505-12. DOI: https://doi.org/10.1016/j.surg.2023.09.046

56. Ashrafi A, Ding L, Atay SM, Wightman SC, Harano T, Kim AW. Delays to surgery and worse outcomes: The compounding effects of social determinants of health in non-small cell lung cancer. *JTCVS Open*. 2023;15:468-78. DOI: 10.1016/j.xjon.2023.05.015

57. Korst MR, Patel AM, Garcia DJ, et al. Disparities in lung cancer short- and long-term outcomes after surgery: Analysis from the national cancer database☆. *Cancer Treatment and Research Communications*. 2023;37:100777. DOI: https://doi.org/10.1016/j.ctarc.2023.100777

58. Duncan FC, Al Nasrallah N, Nephew L, et al. Racial disparities in staging, treatment, and mortality in non-small cell lung cancer. *Transl Lung Cancer Res*. 2024;13(1):76-94. DOI: 10.21037/tlcr-23-407

59. 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*. 2014;83(3):401-7. DOI: https://doi.org/10.1016/j.lungcan.2014.01.008

60. Mehta AJ, Stock S, Gray SW, Nerenz DR, Ayanian JZ, Keating NL. Factors contributing to disparities in mortality among patients with non-small-cell lung cancer. *Cancer Med*. 2018;7(11):5832-42. DOI: 10.1002/cam4.1796

61. Tannenbaum SL, Koru-Sengul T, Zhao W, Miao F, Byrne MM. Survival Disparities in Non–Small Cell Lung Cancer by Race, Ethnicity, and Socioeconomic Status. *The Cancer*

Journal. 2014;20(4). DOI: https://journals.lww.com/journalppo/fulltext/2014/07000/survival_disparities_in_non_small_cell_lung_cancer.1.aspx

62. Zhou K, Shi H, Chen R, et al. Association of Race, Socioeconomic Factors, and Treatment Characteristics With Overall Survival in Patients With Limited-Stage Small Cell Lung Cancer. *JAMA Network Open*. 2021;4(1):e2032276-e. DOI: 10.1001/jamanetworkopen.2020.32276

63. Lin Y, Mak KS. Racial and socioeconomic disparities in the use of stereotactic body radiotherapy for treating non-small cell lung cancer: a narrative review. *J Thorac Dis*. 2021;13(6):3764-71. DOI: 10.21037/jtd-20-3199

64. Lee H, Singh GK. Disparities in All-cancer and Lung Cancer Survival by Social, Behavioral, and Health Status Characteristics in the United States: A Longitudinal Follow-up of the 1997-2015 National Health Interview Survey-National Death Index Record Linkage Study. *J Cancer Prev*. 2022;27(2):89-100. DOI: 10.15430/jcp.2022.27.2.89

65. Harvey-Sullivan A, Ali S, Dhesi P, et al. Comparing cancer stage at diagnosis between migrants and non-migrants: a meta-analysis. *Br J Cancer*. 2024. DOI: 10.1038/s41416-024-02896-0

66. Dalton SO, Steding-Jessen M, Engholm G, Schüz J, Olsen JH. Social inequality and incidence of and survival from lung cancer in a population-based study in Denmark, 1994–2003. *Eur J Cancer*. 2008;44(14):1989-95. DOI: https://doi.org/10.1016/j.ejca.2008.06.023

67. Dalton SO, Steding-Jessen M, Jakobsen E, et al. Socioeconomic position and survival after lung cancer: Influence of stage, treatment and comorbidity among Danish patients with lung cancer diagnosed in 2004–2010. *Acta Oncol*. 2015;54(5):797-804. DOI: 10.3109/0284186X.2014.1001037

68. Pokhrel A, Martikainen P, Pukkala E, Rautalahti M, Seppä K, Hakulinen T. Education, survival and avoidable deaths in cancer patients in Finland. *Br J Cancer*. 2010;103(7):1109-14. DOI: 10.1038/sj.bjc.6605861

69. Kravdal Ø. Social Inequalities in Cancer Survival. *Population Studies*. 2000;54(1):1-18. DOI: http://www.jstor.org/stable/2584629

70. Skyrud KD, Bray F, Eriksen MT, Nilssen Y, Møller B. Regional variations in cancer survival: Impact of tumour stage, socioeconomic status, comorbidity and type of treatment in Norway. *Int J Cancer*. 2016;138(9):2190-200. DOI: https://doi.org/10.1002/ijc.29967

71. Berglund A, Holmberg L, Tishelman C, Wagenius G, Eaker S, Lambe M. Social inequalities in non-small cell lung cancer management and survival: a population-based study in central Sweden. *Thorax*. 2010;65(4):327. DOI: 10.1136/thx.2009.125914

72. Hussain SK, Lenner P, Sundquist J, Hemminki K. Influence of education level on cancer survival in Sweden. *Ann Oncol*. 2008;19(1):156-62. DOI: https://doi.org/10.1093/annonc/mdm413

73. Di Maio M, Signoriello S, Morabito A, et al. Prognostic impact of education level of patients with advanced non-small cell lung cancer enrolled in clinical trials. *Lung Cancer*. 2012;76(3):457-64. DOI: https://doi.org/10.1016/j.lungcan.2012.01.002

74. Pagano E, Filippini C, Di Cuonzo D, et al. Factors affecting pattern of care and survival in a population-based cohort of non-small-cell lung cancer incident cases. *Cancer Epidemiol*. 2010;34(4):483-9. DOI: https://doi.org/10.1016/j.canep.2010.04.002

75. Smailyte G, Jasilionis D, Vincerzevskiene I, Shkolnikov VM. Education, survival, and avoidable deaths in Lithuanian cancer patients, 2001–2009. *Acta Oncol*. 2016;55(7):859-64. DOI: 10.3109/0284186X.2016.1156739

76. Aarts MJ, Kamphuis CBM, Louwman MJ, Coebergh JWW, Mackenbach JP, van Lenthe FJ. Educational inequalities in cancer survival: a role for comorbidities and health behaviours? *J Epidemiol Community Health*. 2013;67(4):365-73. DOI: 10.1136/jech-2012-201404

77. Clément-Duchêne C, Stock S, Xu X, et al. Survival among Never-Smokers with Lung Cancer in the Cancer Care Outcomes Research and Surveillance Study. *Ann Am Thorac Soc*. 2016;13(1):58-66. DOI: 10.1513/AnnalsATS.201504-241OC

78. Herndon JE, 2nd, Kornblith AB, Holland JC, Paskett ED. Patient education level as a predictor of survival in lung cancer clinical trials. *J Clin Oncol*. 2008;26(25):4116-23. DOI: 10.1200/jco.2008.16.7460

79. Yeole BB, Kumar AV. Population-based survival from cancers having a poor prognosis in Mumbai (Bombay), India. *Asian Pac J Cancer Prev*. 2004;5(2):175-82. DOI:

80. Yeole BB. Respiratory cancer population-based survival in Mumbai, India. *Asian Pac J Cancer Prev*. 2005;6(4):449-54. DOI:

81. Fujino Y. Anthropometry, development history and mortality in the Japan Collaborative Cohort Study for Evaluation of Cancer (JACC). *Asian Pac J Cancer Prev*. 2007;8 Suppl:105-12. DOI:

82. Rosso S, Faggiano F, Zanetti R, Costa G. Social class and cancer survival in Turin, Italy. *J Epidemiol Community Health*. 1997;51(1):30-4. DOI: 10.1136/jech.51.1.30

83. Stavraky KM, Skillings JR, Stitt LW, Gwadry-Sridhar F. The effect of socioeconomic status on the long-term outcome of cancer. *J Clin Epidemiol*. 1996;49(10):1155-60. DOI: https://doi.org/10.1016/0895-4356(96)00179-5

84. Malagón T, Botting-Provost S, Moore A, El-Zein M, Franco EL. Inequalities in relative cancer survival by race, immigration status, income, and education for 22 cancer sites in

Canada, a cohort study. *medRxiv*. 2024:2024.05.31.24307976. DOI: 10.1101/2024.05.31.24307976

85. Johnson AM, Johnson A, Hines RB, Bayakly R. The Effects of Residential Segregation and Neighborhood Characteristics on Surgery and Survival in Patients with Early-Stage Non–Small Cell Lung Cancer. *Cancer Epidemiol Biomarkers Prev*. 2016;25(5):750-8. DOI: 10.1158/1055-9965.Epi-15-1126

86. Khullar OV, Gillespie T, Nickleach DC, et al. Socioeconomic Risk Factors for Long-Term Mortality after Pulmonary Resection for Lung Cancer: An Analysis of More than 90,000 Patients from the National Cancer Data Base. *J Am Coll Surg*. 2015;220(2). DOI: https://journals.lww.com/journalacs/fulltext/2015/02000/socioeconomic_risk_factors_for_long_term_mortality.4.aspx

87. Melvan JN, Sancheti MS, Gillespie T, et al. Nonclinical Factors Associated with 30-Day Mortality after Lung Cancer Resection: An Analysis of 215,000 Patients Using the National Cancer Data Base. *J Am Coll Surg*. 2015;221(2). DOI: https://journals.lww.com/journalacs/fulltext/2015/08000/nonclinical_factors_associated_with_30_day.36.aspx

88. Braaten T, Weiderpass E, Lund E. Socioeconomic differences in cancer survival: the Norwegian Women and Cancer Study. *BMC Public Health*. 2009;9:178. DOI: 10.1186/1471-2458-9-178

89. Finke I, Behrens G, Weisser L, Brenner H, Jansen L. Socioeconomic Differences and Lung Cancer Survival-Systematic Review and Meta-Analysis. *Front Oncol*. 2018;8:536-. DOI: https://doi.org/10.3389/fonc.2018.00536

90. Singh GK. Area socioeconomic variations in US cancer incidence, mortality, stage, treatment, and survival, 1975-1999. Bethesda, MD: US Department of Health and Human Services, National Institutes of Health …; 2003.

91. Jeong SM, Jung KW, Park J, Lee HJ, Shin DW, Suh M. Disparities in Overall Survival Rates for Cancers across Income Levels in the Republic of Korea. *Cancers (Basel)*. 2024;16(16). DOI: 10.3390/cancers16162923

92. Exarchakou A, Kipourou DK, Belot A, Rachet B. Socio-economic inequalities in cancer survival: how do they translate into Number of Life-Years Lost? *Br J Cancer*. 2022;126(10):1490-8. DOI: 10.1038/s41416-022-01720-x

93. Habbous S, Khan Y, Langer DL, 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. DOI: https://doi.org/10.4103/atm.atm_283_20

94. Groome PA, Schulze KM, Keller S, Mackillop WJ. Demographic Differences between Cancer Survivors and those who Die Quickly of their Disease. *Clin Oncol*. 2008;20(8):647-56. DOI: https://doi.org/10.1016/j.clon.2008.05.006

95. Shah M, Parmar A, Chan KKW. Socioeconomic disparity trends in diagnostic imaging, treatments, and survival for non-small cell lung cancer 2007-2016. *Cancer medicine*. 2020;9(10):3407-16. DOI: https://doi.org/10.1002/cam4.2978

96. Ruan Y, Heer E, Warkentin MT, et al. The association between neighborhood-level income and cancer stage at diagnosis and survival in Alberta. *Cancer*. 2023. DOI: https://doi.org/10.1002/cncr.35098

97. Canadian Partnership Against Cancer. Lung cancer and equity: A focus on income and geography. Toronto, ON.: Canadian Partnership Against Cancer; 2020.

98. Mullin ML, Shellenberger J, Whitehead M, et al. Characterizing Regional Variability in Lung Cancer Outcomes across Ontario-A Population-Based Analysis. *Curr Oncol*. 2022;29(12):9640-59. DOI: 10.3390/curroncol29120757

99. Redondo-Sánchez D, Petrova D, Rodríguez-Barranco M, Fernández-Navarro P, Jiménez-Moleón JJ, Sánchez MJ. Socio-Economic Inequalities in Lung Cancer Outcomes: An Overview of Systematic Reviews. *Cancers (Basel)*. 2022;14(2). DOI: 10.3390/cancers14020398

100. Mackillop WJ, Zhang-Salomons J, Groome PA, Paszat L, Holowaty E. Socioeconomic status and cancer survival in Ontario. *J Clin Oncol*. 1997;15(4):1680-9. DOI: https://doi.org/10.1200/JCO.1997.15.4.1680

101. Booth CM, Li G, Zhang-Salomons J, Mackillop WJ. The impact of socioeconomic status on stage of cancer at diagnosis and survival. *Cancer*. 2010;116(17):4160-7. DOI: https://doi.org/10.1002/cncr.25427

102. Dabbikeh A, Peng Y, Mackillop WJ, Booth CM, Zhang-Salomons J. Temporal trends in the association between socioeconomic status and cancer survival in Ontario: a population-based retrospective study. *CMAJ Open*. 2017;5(3):E682-e9. DOI: 10.9778/cmajo.20170025

103. 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*. 1999;17(7):2244-. DOI: https://doi.org/10.1200/JCO.1999.17.7.2244

104. Zhang–Salomons J, Qian H, Holowaty E, Mackillop WJ. Associations Between Socioeconomic Status and Cancer Survival: Choice of SES Indicator May Affect Results. *Ann Epidemiol*. 2006;16(7):521-8. DOI: https://doi.org/10.1016/j.annepidem.2005.10.002

105. Hastert TA, Beresford SA, Sheppard L, White E. Disparities in cancer incidence and mortality by area-level socioeconomic status: a multilevel analysis. *J Epidemiol Community Health*. 2015;69(2):168-76. DOI: 10.1136/jech-2014-204417

106. Erhunmwunsee L, Joshi M-BM, Conlon DH, Harpole Jr. DH. Neighborhood-level socioeconomic determinants impact outcomes in nonsmall cell lung cancer patients in the

Southeastern United States. *Cancer*. 2012;118(20):5117-23. DOI: https://doi.org/10.1002/cncr.26185

107. Greenwald HP, Polissar NL, Borgatta EF, McCorkle R. Detecting survival effects of socioeconomic status: Problems in the use of aggregate measures. *J Clin Epidemiol*. 1994;47(8):903-9. DOI: https://doi.org/10.1016/0895-4356(94)90194-5

108. Greenwald HP, Polissar NL, Borgatta EF, McCorkle R, Goodman G. Social factors, treatment, and survival in early-stage non-small cell lung cancer. *Am J Public Health*. 1998;88(11):1681-4. DOI: 10.2105/ajph.88.11.1681

109. Lipworth L, Abelin T, Connelly RR. Socio-economic factors in the prognosis of cancer patients. *J Chronic Dis*. 1970;23(2):105-16. DOI: https://doi.org/10.1016/0021-9681(70)90069-X

110. Shugarman LR, Sorbero ME, Tian H, Jain AK, Ashwood JS. An exploration of urban and rural differences in lung cancer survival among medicare beneficiaries. *Am J Public Health*. 2008;98(7):1280-7. DOI: 10.2105/ajph.2006.099416

111. Bonett A, Roder D, Esterman A. Determinants of case survival for cancers of the lung, colon, breast and cervix in South Australia. *Med J Aust*. 1984;141(11):705-9. DOI: https://doi.org/10.5694/j.1326-5377.1984.tb113226.x

112. Caposole MZ, Miller K, Kim JN, Steward NA, Bauer TL. Elimination of socioeconomic and racial disparities related to lung cancer: closing the gap at a high volume community cancer center. *Surg Oncol*. 2014;23(2):46-52. DOI: 10.1016/j.suronc.2014.02.001

113. Niu X, Pawlish KS, Roche LM. Cancer survival disparities by race/ethnicity and socioeconomic status in New Jersey. *J Health Care Poor Underserved*. 2010;21(1):144-60. DOI: 10.1353/hpu.0.0263

114. Gorey KM, Holowaty EJ, Fehringer G, et al. An international comparison of cancer survival: Toronto, Ontario, and Detroit, Michigan, metropolitan areas. *Am J Public Health*. 1997;87(7):1156-63. DOI: 10.2105/ajph.87.7.1156

115. Yang R, Cheung MC, Byrne MM, et al. Do racial or socioeconomic disparities exist in lung cancer treatment? *Cancer*. 2010;116(10):2437-47. DOI: https://doi.org/10.1002/cncr.24986

116. Wang S, Sun T, Sun H, et al. Survival improvement in patients with non–small cell lung cancer between 1983 and 2012: Analysis of the Surveillance, Epidemiology, and End Results database. *Tumour Biol*. 2017;39(5):1010428317691677. DOI: 10.1177/1010428317691677

117. Wang S, Tang J, Sun T, et al. Survival changes in patients with small cell lung cancer and disparities between different sexes, socioeconomic statuses and ages. *Sci Rep*. 2017;7(1):1339. DOI: 10.1038/s41598-017-01571-0

118. Berglund A, Lambe M, Lüchtenborg M, et al. Social differences in lung cancer management and survival in South East England: a cohort study. *BMJ Open*. 2012;2(3):e001048. DOI: 10.1136/bmjopen-2012-001048

119. Rachet B, Quinn MJ, Cooper N, Coleman MP. Survival from cancer of the lung in England and Wales up to 2001. *Br J Cancer*. 2008;99 Suppl 1(Suppl 1):S40-2. DOI: 10.1038/sj.bjc.6604583

120. Riaz SP, Horton M, Kang J, Mak V, Lüchtenborg M, Møller H. Lung Cancer Incidence and Survival in England: An Analysis by Socioeconomic Deprivation and Urbanization. *J Thorac Oncol*. 2011;6(12):2005-10. DOI: 10.1097/JTO.0b013e31822b02db

121. Forrest LF, Adams J, Rubin G, White M. The role of receipt and timeliness of treatment in socioeconomic inequalities in lung cancer survival: population-based, data-linkage study. *Thorax*. 2015;70(2):138-45. DOI: 10.1136/thoraxjnl-2014-205517

122. Doiron D, Setton EM, Shairsingh K, et al. Healthy built environment: Spatial patterns and relationships of multiple exposures and deprivation in Toronto, Montreal and Vancouver. *Environ Int*. 2020;143:106003. DOI: https://doi.org/10.1016/j.envint.2020.106003

123. Eagles ME, Beall RF, Ben-Israel D, Wong JH, Hill MD, Spackman E. Neighbourhood deprivation, distance to nearest comprehensive stroke centre and access to endovascular thrombectomy for ischemic stroke: a population-based study. *CMAJ Open*. 2023;11(6):E1181-E7. DOI: 10.9778/cmajo.20230046

124. Messer LC, Laraia BA, Kaufman JS, et al. The Development of a Standardized Neighborhood Deprivation Index. *J Urban Health*. 2006;83(6):1041-62. DOI: 10.1007/s11524-006-9094-x

125. Harding J, Stevenson C, Sadkowsky K, Markey J, Leeds K, editors. Cancer survival in Australia 1992–1997 : geographic categories and socioeconomic status2003.

126. Thomas AA, Pearce A, O'Neill C, Molcho M, Sharp L. Urban-rural differences in cancer-directed surgery and survival of patients with non-small cell lung cancer. *J Epidemiol Community Health*. 2017;71(5):468-74. DOI: 10.1136/jech-2016-208113

127. Finke I, Behrens G, Schwettmann L, et al. Socioeconomic differences and lung cancer survival in Germany: Investigation based on population-based clinical cancer registration. *Lung Cancer*. 2020;142:1-8. DOI: https://doi.org/10.1016/j.lungcan.2020.01.021

128. Pilleron S, Maringe C, Charvat H, Atkinson J, Morris E, Sarfati D. Age disparities in lung cancer survival in New Zealand: The role of patient and clinical factors. *Lung Cancer*. 2021;157:92-9. DOI: https://doi.org/10.1016/j.lungcan.2021.05.015

129. Zadnik V, Žagar T, Tomšič S, Mihor A, Lokar K. Cancer Patients' Survival According to Socioeconomic Environment in a High-Income Country with Universal Health Coverage. *Cancers (Basel)*. 2022;14(7). DOI: 10.3390/cancers14071620

130. Rahman MM, David M, Goldsbury D, et al. Impact of cancer diagnosis on life expectancy by area-level socioeconomic groups in New South Wales, Australia: a population-based study. *Cancer Biol Med*. 2024;21(8):692-702. DOI: 10.20892/j.issn.2095-3941.2024.0166

131. Lennox L, Lambe K, Hindocha CN, Coronini-Cronberg S. What health inequalities exist in access to, outcomes from and experience of treatment for lung cancer? A scoping review. *BMJ Open*. 2023;13(11):e077610. DOI: 10.1136/bmjopen-2023-077610

132. May L, Shows K, Nana-Sinkam P, Li H, Landry JW. Sex Differences in Lung Cancer. *Cancers (Basel)*. 2023;15(12):3111. DOI: https://www.mdpi.com/2072-6694/15/12/3111

133. Pal SK, Hurria A. Impact of age, sex, and comorbidity on cancer therapy and disease progression. *J Clin Oncol*. 2010;28(26):4086-93. DOI: 10.1200/jco.2009.27.0579

134. World Health Organization. *Global report on ageism*: World Health Organization; 2021.

135. Chang ES, Kannoth S, Levy S, Wang SY, Lee JE, Levy BR. Global reach of ageism on older persons' health: A systematic review. *PLoS One*. 2020;15(1):e0220857. DOI: 10.1371/journal.pone.0220857

136. Federal/Provincial and Territorial (FPT) Ministers Responsible for Seniors. Addressing the social and economic impacts of ageism in Canada. A discussion guide on ageism in Canada. 2022.

137. Walker MJ, Meggetto O, Gao J, et al. Measuring the impact of the COVID-19 pandemic on organized cancer screening and diagnostic follow-up care in Ontario, Canada: A provincial, population-based study. *Prev Med*. 2021;151:106586. DOI: 10.1016/j.ypmed.2021.106586

138. Ramanakumar AV, Annie B, Frederic L, Christine B, Cathy R, Jean L. Evaluating the impact of COVID-19 on cancer declarations in Quebec, Canada. *Cancer Med*. 2023;12(5):6260-9. DOI: 10.1002/cam4.5389

139. Kim JO, Davis F, Butts C, Winget M. Waiting Time Intervals for Non-small Cell Lung Cancer Diagnosis and Treatment in Alberta: Quantification of Intervals and Identification of Risk Factors Associated with Delays. *Clin Oncol (R Coll Radiol)*. 2016;28(12):750-9. DOI: 10.1016/j.clon.2016.06.010

140. Sacher AG, Le LW, Leighl NB, Coate LE. Elderly Patients with Advanced NSCLC in Phase III Clinical Trials: Are the Elderly Excluded from Practice-Changing Trials in Advanced NSCLC? *J Thorac Oncol*. 2013;8(3):366-8. DOI: 10.1097/JTO.0b013e31827e2145

141. Yee KW, Pater JL, Pho L, Zee B, Siu LL. Enrollment of older patients in cancer treatment trials in Canada: why is age a barrier? *J Clin Oncol*. 2003;21(8):1618-23. DOI: 10.1200/jco.2003.12.044

142. Javier-DesLoges J, Nelson TJ, Murphy JD, et al. Disparities and trends in the participation of minorities, women, and the elderly in breast, colorectal, lung, and prostate cancer clinical trials. *Cancer*. 2022;128(4):770-7. DOI: https://doi.org/10.1002/cncr.33991

143. Haase KR, Sattar S, Pilleron S, et al. A scoping review of ageism towards older adults in cancer care. *J Geriatr Oncol*. 2023;14(1):101385. DOI: https://doi.org/10.1016/j.jgo.2022.09.014

144. Cook S, Alibhai S, Mehta R, et al. Improving Care for Older Adults with Cancer in Canada: A Call to Action. *Current Oncology*. 2024;31(7):3783-97. DOI: https://www.mdpi.com/1718-7729/31/7/279

145. Florez N, Kiel L, Riano I, et al. Lung Cancer in Women: The Past, Present, and Future. *Clin Lung Cancer*. 2024;25(1):1-8. DOI: https://doi.org/10.1016/j.cllc.2023.10.007

146. Monestime S, Rigney M, Phipps N, et al. Health inequities across the lung cancer care continuum in ten marginalized populations: a narrative review. *J Thorac Dis*. 2023;15(11):6345-61. DOI: 10.21037/jtd-23-727

147. Ragavan M, Patel MI. The evolving landscape of sex-based differences in lung cancer: a distinct disease in women. *European Respiratory Review*. 2022;31(163):210100. DOI: 10.1183/16000617.0100-2021

148. Gee K, Yendamuri S. Lung cancer in females-sex-based differences from males in epidemiology, biology, and outcomes: a narrative review. *Transl Lung Cancer Res*. 2024;13(1):163-78. DOI: 10.21037/tlcr-23-744

149. Tolwin Y, Gillis R, Peled N. Gender and lung cancer—SEER-based analysis. *Ann Epidemiol*. 2020;46:14-9. DOI: https://doi.org/10.1016/j.annepidem.2020.04.003

150. Araghi M, Fidler-Benaoudia M, Arnold M, et al. International differences in lung cancer survival by sex, histological type and stage at diagnosis: an ICBP SURVMARK-2 Study. Thorax [Internet]. 2022; 77(4):[378-90 pp.]. Available from: http://dx.doi.org/10.1136/thoraxjnl-2020-216555.

151. Rana RH, Alam F, Alam K, Gow J. Gender-specific differences in care-seeking behaviour among lung cancer patients: a systematic review. *J Cancer Res Clin Oncol*. 2020;146(5):1169-96. DOI: 10.1007/s00432-020-03197-8

152. Pilleron S, Gower H, Janssen-Heijnen M, et al. Patterns of age disparities in colon and lung cancer survival: a systematic narrative literature review. *BMJ Open*. 2021;11(3):e044239. DOI: 10.1136/bmjopen-2020-044239

153. Ahmed S, Shahid RK. Disparity in cancer care: a Canadian perspective. *Curr Oncol*. 2012;19(6):e376-e82. DOI: 10.3747/co.19.1177

154. Maddison AR, Asada Y, Urquhart R. Inequity in access to cancer care: a review of the Canadian literature. *Cancer Causes Control*. 2011;22(3):359-66. DOI: 10.1007/s10552-010-9722-3

155. Kobayashi LC, Westrick AC, Doshi A, et al. New directions in cancer and aging: State of the science and recommendations to improve the quality of evidence on the intersection of aging with cancer control. *Cancer*. 2022;128(9):1730-7. DOI: 10.1002/cncr.34143

156. Hankivsky O, Jordan-Zachery JS. *The Palgrave Handbook of Intersectionality in Public Policy*: Springer International Publishing; 2019. Available from: https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5667407.

157. Statistics Canada. Immigrants make up the largest share of the population in over 150 years and continue to shape who we are as Canadians. *The Daily*. 2022:1-21. DOI:

158. Chui T, Statistics C. *Immigration and ethnocultural diversity in Canada : National Household Survey, 2011*. [Ottawa]: Statistics Canada = Statistique Canada; 2013. Available from: https://central.bac-lac.gc.ca/.item?id=CS99-010-2011-1-eng&op=pdf&app=Library.

159. Lu C, Ng E. Healthy immigrant effect by immigrant category in Canada. Health Rep: Statistics Canada; 2019. p. 3-11.

160. Vang Z, Sigouin J, Flenon A, Gagnon A. The healthy immigrant effect in Canada: A systematic review. *Population Change and Lifecourse Strategic Knowledge Cluster Discussion Paper Series/Un Réseau stratégique de connaissances Changements de population et parcours de vie Document de travail*. 2015;3(1):4. DOI:

161. Mikkonen J, Raphael D, York University . School of Health P, Management. *Social determinants of health : the Canadian facts*. [Toronto]: [York University, School of Health Policy and Management]; 2010. Available from: https://www.deslibris.ca/ID/223317.

162. Kalich A, Heinemann L, Ghahari S. A Scoping Review of Immigrant Experience of Health Care Access Barriers in Canada. *Journal of Immigrant and Minority Health*. 2016;18(3):697-709. DOI: 10.1007/s10903-015-0237-6

163. Gomez SL, Clarke CA, Shema SJ, Chang ET, Keegan THM, Glaser SL. Disparities in Breast Cancer Survival Among Asian Women by Ethnicity and Immigrant Status: A Population-Based Study. *Am J Public Health*. 2010;100(5):861-9. DOI: 10.2105/ajph.2009.176651

164. Montealegre JR, Zhou R, Amirian ES, Follen M, Scheurer ME. Nativity disparities in late-stage diagnosis and cause-specific survival among Hispanic women with invasive cervical cancer: an analysis of Surveillance, Epidemiology, and End Results data. *Cancer Causes Control*. 2013;24(11):1985-94. DOI: 10.1007/s10552-013-0274-1

165. Agaku IT, Adisa AO. Nativity status and oral cancer survival in the United States: Implications for dental clinical practice. *Quintessence Int*. 2014;45(4). DOI:

166. Santos PMG, Boe LA, Barnes JM, Cervantes L, Yabroff R, Chino F. Disparities in cancer mortality outcomes among adult immigrants in the United States. American Society of Clinical Oncology; 2024.

167. Shariff-Marco S, Gomez SL, Canchola AJ, et al. Nativity, ethnic enclave residence, and breast cancer survival among Latinas: Variations between California and Texas. *Cancer*. 2020;126(12):2849-58. DOI: https://doi.org/10.1002/cncr.32845

168. Zhou K, Song Z, Rostomian N, et al. Association of nativity with survival among adults with hepatocellular carcinoma. *JNCI: Journal of the National Cancer Institute*. 2023;115(7):861-9. DOI: 10.1093/jnci/djad067

169. Schupp CW, Press DJ, Gomez SL. Immigration factors and prostate cancer survival among Hispanic men in California: Does neighborhood matter? *Cancer*. 2014;120(9):1401-8. DOI: https://doi.org/10.1002/cncr.28587

170. Freedland AR, Sanchez Mendez J, Liu L, et al. Prostate Cancer Disparities in Clinical Characteristics and Survival among Black and Latino Patients Considering Nativity: Findings from the California Cancer Registry. *Cancer Epidemiol Biomarkers Prev*. 2024;33(11):1512-22. DOI: 10.1158/1055-9965.Epi-24-0678

171. Keegan THM, Quach T, Shema S, Glaser SL, Gomez SL. The influence of nativity and neighborhoods on breast cancer stage at diagnosis and survival among California Hispanic women. *BMC Cancer*. 2010;10(1):603. DOI: 10.1186/1471-2407-10-603

172. Hagopian GS, Bapat AC, Dottino PR, et al. The impact of nativity on uterine cancer survival in the public hospital system of Queens, New York. *Gynecol Oncol*. 2024;181:133-40. DOI: https://doi.org/10.1016/j.ygyno.2023.12.018

173. Chau B, Ituarte PH, Shinde A, et al. Disparate outcomes in nonsmall cell lung cancer by immigration status. *Cancer Med*. 2021. DOI: 10.1002/cam4.3848

174. Cheung MC, Earle CC, Fischer HD, et al. Impact of Immigration Status on Cancer Outcomes in Ontario, Canada. *J Oncol Pract*. 2017;13(7):e602-e12. DOI: https://doi.org/10.1200/JOP.2016.019497

175. Ruco A, Lofters AK, Lu H, et al. Lung cancer survival by immigrant status: a population-based retrospective cohort study in Ontario, Canada. *BMC Cancer*. 2024;24(1):1114. DOI: 10.1186/s12885-024-12804-7

176. Zarifa D, Seward B, Milian RP. Location, location, location: Examining the rural-urban skills gap in Canada. *Journal of Rural Studies*. 2019;72:252-63. DOI: https://doi.org/10.1016/j.jrurstud.2019.10.032

177. Fayet Y, Praud D, Fervers B, et al. Beyond the map: evidencing the spatial dimension of health inequalities. *International Journal of Health Geographics*. 2020;19(1):46. DOI: 10.1186/s12942-020-00242-0

178. Kondro W. Health of rural Canadians lags behind urban counterparts. *CMAJ*. 2006;175(10):1195. DOI: 10.1503/cmaj.061306

179. Mitura V, Bollman RD. The health of rural Canadians: A rural-urban comparison of health indicators. In: Canada S, editor. 2003.

180. Canadian Institute for Health Information. How healthy are rural Canadians? An assessment of their health status and health determinants. A component of the initiative "Canada's Rural Communities : understanding rural health and its determinants.". In: Canadian Population Health Initiative, editor. Ottawa, Ont.: Canadian Institute for Health Information; 2006.

181. Carriere R, Adam R, Fielding S, Barlas R, Ong Y, Murchie P. Rural dwellers are less likely to survive cancer – An international review and meta-analysis. *Health & Place*. 2018;53:219-27. DOI: https://doi.org/10.1016/j.healthplace.2018.08.010

182. Canadian Partnership Against Cancer. Examining disparities in cancer control: A system performance special focus report. Canadian Partnership Against Cancer; 2014.

183. Shah BD, Tyan C-C, Rana M, et al. Rural vs urban inequalities in stage at diagnosis for lung cancer. *Cancer Treatment and Research Communications*. 2021;29:100495. DOI: https://doi.org/10.1016/j.ctarc.2021.100495

184. Lambert LK, Horrill TC, Beck SM, et al. Health and healthcare equity within the Canadian cancer care sector: a rapid scoping review. *Int J Equity Health*. 2023;22(1):20. DOI: 10.1186/s12939-023-01829-2

185. Afshar N, English DR, Milne RL. Factors Explaining Socio-Economic Inequalities in Cancer Survival: A Systematic Review. *Cancer Control*. 2021;28:10732748211011956. DOI: https://doi.org/10.1177/10732748211011956

186. O'Rourke N, Edwards R. Lung Cancer Treatment Waiting Times and Tumour Growth. *Clin Oncol*. 2000;12(3):141-4. DOI: https://doi.org/10.1053/clon.2000.9139

187. Ellis PM, Vandermeer R. Delays in the diagnosis of lung cancer. *J Thorac Dis*. 2011;3(3):183-8. DOI: 10.3978/j.issn.2072-1439.2011.01.01

188. Ali N, Lifford KJ, Carter B, et al. Barriers to uptake among high-risk individuals declining participation in lung cancer screening: a mixed methods analysis of the UK Lung Cancer Screening (UKLS) trial. *BMJ Open*. 2015;5(7):e008254. DOI: 10.1136/bmjopen-2015-008254

189. Cao DJ, Alabousi M, Farshait N, Patlas MN. Barriers to Screening At-risk Populations in Canada. *Can Assoc Radiol J*. 2023;74(3):508-13. DOI: 10.1177/08465371221147307

190. Lofters AK, Mark A, Taljaard M, Green ME, Glazier RH, Dahrouge S. Cancer screening inequities in a time of primary care reform: a population-based longitudinal study in Ontario, Canada. *BMC Fam Pract*. 2018;19(1):147. DOI: 10.1186/s12875-018-0827-1

191. Adefemi K. Evaluating the uptake and disparities in colorectal cancer screening in Canada: a cross-sectional study using data from the Canadian Community Health Surveys: Memorial University of Newfoundland; 2024.

192. Ramji F, Cotterchio M, Manno M, Rabeneck L, Gallinger S. Association between subject factors and colorectal cancer screening participation in Ontario, Canada. *Cancer Detect Prev*. 2005;29(3):221-6. DOI: 10.1016/j.cdp.2005.04.001

193. Sayani A, Vahabi M, O'Brien MA, et al. Perspectives of family physicians towards access to lung cancer screening for individuals living with low income - a qualitative study. *BMC Fam Pract*. 2021;22(1):10. DOI: 10.1186/s12875-020-01354-z

194. Ching-López A, Luque-Fernandez MA, Sánchez MJ. Social Disparities in Survival from Lung Cancer in Europe. In: Launoy G, Zadnik V, Coleman MP, editors. *Social Environment and Cancer in Europe: Towards an Evidence-Based Public Health Policy*. Cham: Springer International Publishing; 2021. p. 121-40.

195. Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer survival: a review. *Ann Oncol*. 2006;17(1):5-19. DOI: https://doi.org/10.1093/annonc/mdj007

196. Forrest LF, Sowden S, Rubin G, White M, Adams J. Socio-economic inequalities in stage at diagnosis, and in time intervals on the lung cancer pathway from first symptom to treatment: systematic review and meta-analysis. *Thorax*. 2017;72(5):430. DOI: https://doi.org/10.1136/thoraxjnl-2016-209013

197. Verplancke KB, Keirns DL, McMahon K, et al. Association Between Demographic and Socioeconomic Factors and Diagnosis of Advanced Non-small Cell Lung Cancer: An Analysis of the National Cancer Database. *Cureus*. 2023;15(8):e44351. DOI: 10.7759/cureus.44351

198. Lofters AK, Gatov E, Lu H, et al. Lung Cancer Inequalities in Stage of Diagnosis in Ontario, Canada. *Current Oncology*. 2021;28(3):1946-56. DOI: https://doi.org/10.3390%2Fcurroncol28030181

199. O'Rourke MA, Feussner JR, Feigl P, Laszlo J. Age trends of lung cancer stage at diagnosis. Implications for lung cancer screening in the elderly. *JAMA*. 1987;258(7):921-6. DOI:

200. Subramanian J, Morgensztern D, Goodgame B, et al. Distinctive Characteristics of Non-small Cell Lung Cancer (NSCLC) in the Young: A Surveillance, Epidemiology, and End Results (SEER) Analysis. *J Thorac Oncol*. 2010;5(1):23-8. DOI: https://doi.org/10.1097/JTO.0b013e3181c41e8d

201. Ganti AK, Klein AB, Cotarla I, Seal B, Chou E. Update of Incidence, Prevalence, Survival, and Initial Treatment in Patients With Non–Small Cell Lung Cancer in the US. *JAMA Oncology*. 2021;7(12):1824-32. DOI: 10.1001/jamaoncol.2021.4932

202. Afshar N, English DR, Milne RL. Rural–urban residence and cancer survival in high-income countries: A systematic review. *Cancer*. 2019;125(13):2172-84. DOI: https://doi.org/10.1002/cncr.32073

203. Forrest LF, Adams J, Wareham H, Rubin G, White M. Socioeconomic inequalities in lung cancer treatment: systematic review and meta-analysis. *PLoS Med*. 2013;10(2):e1001376-e. DOI: 10.1371/journal.pmed.1001376

204. Ellis PM, Verma S, Sehdev S, Younus J, Leighl NB. Challenges to implementation of an epidermal growth factor receptor testing strategy for non-small-cell lung cancer in a publicly funded health care system. *J Thorac Oncol*. 2013;8(9):1136-41. DOI: 10.1097/JTO.0b013e31829f6a43

205. Brenner DR, O'Sullivan DE, Jarada TN, et al. The impact of population-based EGFR testing in non-squamous metastatic non-small cell lung cancer in Alberta, Canada. *Lung Cancer*. 2023;175:60-7. DOI: https://doi.org/10.1016/j.lungcan.2022.11.017

206. O'Sullivan DE, Jarada TN, Yusuf A, et al. Prevalence, Treatment Patterns, and Outcomes of Individuals with EGFR Positive Metastatic Non-Small Cell Lung Cancer in a Canadian Real-World Setting: A Comparison of Exon 19 Deletion, L858R, and Exon 20 Insertion EGFR Mutation Carriers. Current Oncology [Internet]. 2022; 29(10):[7198-208 pp.].

207. Perdrizet K, Sutradhar R, Li Q, Liu N, Earle CC, Leighl NB. Second and later-line erlotinib use in non-small cell lung cancer: real world outcomes and practice patterns overtime in Canada. *J Thorac Dis*. 2021;13(9):5419-29. DOI: 10.21037/jtd-21-804

208. Cheema PK, Gomes M, Banerji S, et al. Consensus recommendations for optimizing biomarker testing to identify and treat advanced EGFR-mutated non-small-cell lung cancer. *Curr Oncol*. 2020;27(6):321-9. DOI: 10.3747/co.27.7297

209. Snow S, Brezden-Masley C, Carter MD, et al. Barriers and Unequal Access to Timely Molecular Testing Results: Addressing the Inequities in Cancer Care Delays across Canada. *Current Oncology*. 2024;31(3):1359-75. DOI: https://www.mdpi.com/1718-7729/31/3/103

210. MacPhail C, Snow S. Not All Canadian Cancer Patients Are Equal-Disparities in Public Cancer Drug Funding across Canada. *Curr Oncol*. 2022;29(3):2064-72. DOI: 10.3390/curroncol29030166

211. Patented Medicine Prices Review Board. Alignment Among Public Formularies in Canada, Part 2: Oncology Medicines. Ottawa, ON: Patented Medicine Prices Review Board,; 2021.

212. Lamb-Palmer D, Loschmann C, Henricks P, Shen J, Downson J, Mohideen S. Uncovering the Hidden Costs of Take-Home Cancer Drugs. PDCI Market Access, McKesson Canada; 2021.

213. Morgan SG, Gagnon MA, Charbonneau M, Vadeboncoeur A. Evaluating the effects of Quebec's private-public drug insurance system. *CMAJ*. 2017;189(40):E1259-e63. DOI: 10.1503/cmaj.170726

214. Rebić N, Cheng L, Law MR, Cragg JJ, Brotto LA, De Vera MA. Predictors of cost-related medication nonadherence in Canada: a repeated cross-sectional analysis of the Canadian Community Health Survey. *Can Med Assoc J*. 2024;196(40):E1331-E40. DOI: 10.1503/cmaj.241024

215. Hughes D. L'accès aux médicaments anticancéreux onéreux au Canada : une comparaison entre le Québec, l'Ontario et la Colombie-Britannique. *Pratiques et Organisation des Soins*. 2012;43:9. DOI: 10.3917/pos.431.0009

216. Sorin M, Franco EL, Quesnel-Vallée A. Inter- and intraprovincial inequities in public coverage of cancer drug programs across Canada: a plea for the establishment of a pan-Canadian pharmacare program. *Curr Oncol*. 2019;26(4):266-9. DOI: 10.3747/co.26.4867

217. Lim C, Tsao MS, Le LW, et al. Biomarker testing and time to treatment decision in patients with advanced nonsmall-cell lung cancer†. *Ann Oncol*. 2015;26(7):1415-21. DOI: https://doi.org/10.1093/annonc/mdv208

218. Public Health Agency of Canada. *How to integrate intersectionality theory in quantitative health equity analysis? : a rapid review and checklist of promising practices*. Ottawa, ON: Public Health Agency of Canada = Agence de la santé publique du Canada; 2022. Available from: http://central.bac-lac.gc.ca/.redirect?app=damspub&id=8d9b6779-4cdf-44bb-aa37-72c3c6c5857e.

219. 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. *International Journal for Equity in Health*. 2023;22(1):58. DOI: 10.1186/s12939-023-01860-3

220. Varcoe C, Habib S, Sinding C, Fitch M, Magee Gullatte M, Cummings GG. Health disparities in cancer care: Exploring Canadian, American and international perspectives. *2015*. 2015;25(1):9. DOI: DOI: 10.5737/236880762517381

221. Canadian Cancer Society. Advancing health equity through cancer information and support services: report on communities that are underserved. 2023.

222. Horrill TC, Browne AJ, Stajduhar KI. Equity-Oriented Healthcare: What It Is and Why We Need It in Oncology. *Curr Oncol*. 2022;29(1):186-92. DOI: 10.3390/curroncol29010018

223. Horrill TC, Bourgeois A, Kleijberg M, Linton J, Leahy K, Stajduhar KI. Services, models of care, and interventions to improve access to cancer treatment for adults who are socially disadvantaged: A scoping review protocol. *PLoS One*. 2024;19(2):e0296658. DOI: 10.1371/journal.pone.0296658

224. Gupta A, Omeogu CH, Islam JY, Joshi AR, Akinyemiju TF. Association of area-level socioeconomic status and non–small cell lung cancer stage by race/ethnicity and health care–level factors: Analysis of the National Cancer Database. *Cancer*. 2022;128(16):3099-108. DOI: https://doi.org/10.1002/cncr.34327

225. Quesnel-Vallée A, Arpin E. (Re)examining the Canadian Health Care System through an Equity Lens. *Canadian Public Policy*. 2024;50(S1):127-42. DOI: 10.3138/cpp.2023-060

226. Chen JC, Pawlik T, Kelly EP, Obeng-Gyasi S. Intersectionality in patients with cancer: who should care and why? *Future Oncology*. 2022;18(38):4137-40. DOI: 10.2217/fon-2022-0992

227. Hou SHJ, Petrella A, Tulk J, et al. An Evaluation of Racial and Ethnic Representation in Research Conducted with Young Adults Diagnosed with Cancer: Challenges and Considerations for Building More Equitable and Inclusive Research Practices. *Current Oncology*. 2024;31(4):2244-59. DOI: https://www.mdpi.com/1718-7729/31/4/166

228. Estupiñán Fdez de Mesa M, Marcu A, Ream E, Whitaker KL. Relationship between intersectionality and cancer inequalities: a scoping review protocol. *BMJ Open*. 2023;13(1):e066637. DOI: 10.1136/bmjopen-2022-066637

229. Statistics Canada. Canadian Population Health Survey data (CCHS Annual and Focus Content) integrated with mortality, hospitalization, historical postal codes, cancer, tax data and Census 2020 [updated 2020-06-24. Available from: https://www.statcan.gc.ca/eng/rdc/data/cchs.

230. Postal Code OM Conversion File Plus (PCCF+) [Internet]. Government of Canada. 2023 [cited 2023-12-06]. Available from: https://www150.statcan.gc.ca/n1/en/catalogue/82F0086X#wb-auto-2.

231. Wilkins R. PCCF+ Version 5E User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through March 2009. In: Health Analysis Division, editor. Ottawa, ON.2009.

232. Boyko E, Watkins W. *The Canadian Data Liberation Initiative: An Idea Worth Considering?*: International Household Survey Network; 2011.

233. Institut national d'excellence en santé et en services sociaux (INESSS). Rapport d'activités scientifiques 2016/2017. Québec, Qc: Bibliothèque et Archives nationales du Québec; 2017.

234. Institut national d'excellence en santé et en services sociaux (INESSS). Création et caractérisation d'une cohorte québécoise de patients atteints d'un cancer du poumon à l'aide de données clinico-administratives. Québec, Qc: Bibliothèque et Archives nationales du Québec; 2021. p. 152.

235. Institut national d'excellence en santé et en services sociaux (INESSS). Utilisation en contexte québécois des inhibiteurs de la tyrosine kinase du récepteur du facteur de

croissance épidermique (EGFR) pour le traitement du cancer du poumon. État des pratiques rédigé par Gino Boily, Kossi Thomas Golo, Aude- Christine Guédon, Camille Lehuédé, Samia Qureshi, Ferdaous Roussafi et Erin Strumpf. Québec, Qc2022. p. 109.

236. Canadian Institute for Health Information. Canadian coding standards for version 2022 ICD-10-CA and CCI. CIHI Ottawa, ON; 2022.

237. Statistics Canada. Canadian classification of diagnostic, therapeutic, and surgical procedures. In: Nosology Reference Centre, editor.: Minister of Supply and Services Canada; 1986.

238. Ruano-Raviña A, Provencio M, Calvo de Juan V, et al. Lung cancer symptoms at diagnosis: results of a nationwide registry study. *ESMO Open*. 2020;5(6):e001021. DOI: https://doi.org/10.1136/esmoopen-2020-001021

239. Walter FM, Rubin G, Bankhead C, et al. Symptoms and other factors associated with time to diagnosis and stage of lung cancer: a prospective cohort study. *Br J Cancer*. 2015;112(1):S6-S13. DOI: 10.1038/bjc.2015.30

240. Khare SR, Madathil SA, Batist G, Peter Brojde Lung Cancer G, Vedel I. Lung Cancer Pre-Diagnostic Pathways from First Presentation to Specialist Referral. Current Oncology [Internet]. 2021; 28(1):[378-89 pp.].

241. Kang S, McLeod SL, Walsh C, Grewal K. Patient outcomes associated with cancer diagnosis through the emergency department: A systematic review. *Acad Emerg Med*. 2023;30(9):955-62. DOI: https://doi.org/10.1111/acem.14671

242. Sibley LM, Weiner JP. An evaluation of access to health care services along the rural-urban continuum in Canada. *BMC Health Serv Res*. 2011;11(1):1-11. DOI: https://doi.org/10.1186/1472-6963-11-20

243. McDonald JT, Conde H. Does Geography Matter? The Health Service Use and Unmet Health Care Needs of Older Canadians. *Canadian Journal on Aging / La Revue canadienne du vieillissement*. 2010;29(1):23-37. DOI: 10.1017/S0714980809990389

244. Harrington DW, Wilson K, Rosenberg M, Bell S. Access granted! barriers endure: determinants of difficulties accessing specialist care when required in Ontario, Canada. *BMC Health Serv Res*. 2013;13(1):146. DOI: https://doi.org/10.1186/1472-6963-13-146

245. Demeter SJ, Chmielowiec C, Logus W, et al. The Descriptive Epidemiology of Primary Lung Cancer in an Alberta Cohort with a Mutivariate Analysis of Survival to Two Years. *Can Respir J*. 2003;10(8):729635. DOI: https://doi.org/10.1155/2003/729635

246. Andalib A, Ramana-Kumar AV, Bartlett G, Franco EL, Ferri LE. Influence of Postoperative Infectious Complications on Long-Term Survival of Lung Cancer Patients: A Population-Based Cohort Study. *J Thorac Oncol*. 2013;8(5):554-61. DOI: https://doi.org/10.1097/JTO.0b013e3182862e7e

247. Ashworth A, Kong W, Owen T, Hanna TP, Brundage M. The management of stage I Non-Small cell lung cancer (NSCLC) in Ontario: A Population-Based study of patterns of care and Stereotactic Ablative Body radiotherapy (SABR) utilization from 2010 to 2019. *Radiother Oncol*. 2024;194:110153. DOI: https://doi.org/10.1016/j.radonc.2024.110153

248. Butler SJ, Louie AV, Sutradhar R, Paszat L, Brooks D, Gershon AS. Association between COPD and Stage of Lung Cancer Diagnosis: A Population-Based Study. *Current Oncology*. 2023;30(7):6397-410. DOI: https://www.mdpi.com/1718-7729/30/7/471

249. Finley C, Begum H, Akhtar-Danesh G-G, Akhtar-Danesh N. Survival effects of time to surgery for Stage I lung cancer: A population-based study. *Surg Oncol*. 2022;42:101744. DOI: https://doi.org/10.1016/j.suronc.2022.101744

250. McFadgen H, Lethbridge L, Dunbar M, Plourde M. Trends in lung cancer diagnosis and care among visible minorities and non-visible minorities in Canada. *Cancer Causes Control*. 2024;35(4):575-82. DOI: 10.1007/s10552-023-01816-6

251. Liu HW, Kerba M, Lim G, et al. Factors Associated With the Use of Radiation Therapy in Patients With Stage III Non-small Cell Lung Cancer in Alberta, Canada: A Population-based Study. *Cureus*. 2016;8(10):e851. DOI: 10.7759/cureus.851

252. Ellis PM, Verma S, Sehdev S, Younus J, Leighl NB. Challenges to Implementation of an Epidermal Growth Factor Receptor Testing Strategy for Non–Small-Cell Lung Cancer in a Publicly Funded Health Care System. *J Thorac Oncol*. 2013;8(9):1136-41. DOI: https://doi.org/10.1097/JTO.0b013e31829f6a43

253. Blanc-Durand F, Florescu M, Tehfe M, et al. Improvement of EGFR Testing over the Last Decade and Impact of Delaying TKI Initiation. *Curr Oncol*. 2021;28(2):1045-55. DOI: 10.3390/curroncol28020102

254. Myles J, Hou F. Neighbourhood attainment and residential segregation among Toronto's visible minorities. *Statistics Canada, Analytical Studies Branch Research Paper Series*. 2003(206). DOI:

255. Moroz I, D'Angelo MS. A Blueprint for Optimizing Lung Cancer Care for Better Patient Outcomes. Ottawa: The Conference Board of Canada; 2021.

256. Bédard S, Sasewich H, Culling J, et al. Stigma in Early-Stage Lung Cancer. *Ann Behav Med*. 2022;56(12):1272-83. DOI: 10.1093/abm/kaac021

257. Sayani A, Vahabi M, O'Brien MA, et al. Advancing health equity in cancer care: The lived experiences of poverty and access to lung cancer screening. *PLoS One*. 2021;16(5):e0251264. DOI: 10.1371/journal.pone.0251264

258. Torchalla I, Strehlau V, Okoli CTC, Li K, Schuetz C, Krausz M. Smoking and Predictors of Nicotine Dependence in a Homeless Population. *Nicotine & Tobacco Research*. 2011;13(10):934-42. DOI: 10.1093/ntr/ntr101

259. Cowan L, Hwang S, Khandor E, Mason K. The street health report 2007: Highlights & Action Plan. Street Health; 2007.

260. Baggett TP, Hwang SW, O'Connell JJ, et al. Mortality Among Homeless Adults in Boston: Shifts in Causes of Death Over a 15-Year Period. *JAMA Internal Medicine*. 2013;173(3):189-95. DOI: 10.1001/jamainternmed.2013.1604

261. Quayum S, Love C, Hunter P. Everyone Counts 2020-2022 – Results from the Third Nationally Coordinated Point-in- Time Counts of Homelessness in Canada. In: Infrastructure Canada, editor. Ottawa, ON: Minister of Housing, Infrastructure and Communities; 2024.

262. Alston J, Baral S, Orkin A, Straus S. Tackling late-life homelessness in Canada. *Can Med Assoc J*. 2024;196(19):E662-E5. DOI: 10.1503/cmaj.231493

263. O'Grady JR, Ferdus J, Leylachian S, et al. Lung cancer in First Nations, Inuit, and Métis peoples in Canada - a scoping review. *Int J Circumpolar Health*. 2024;83(1):2381879. DOI: 10.1080/22423982.2024.2381879

264. Klugman M, Xue X, Hosgood HD. Race/ethnicity and lung cancer survival in the United States: a meta-analysis. *Cancer Causes Control*. 2019;30(11):1231-41. DOI: 10.1007/s10552-019-01229-4

265. Institut national de santé publique du Québec. *Portrait de santé du Québec et de ses régions 2006: deuxième rapport national sur l'état de santé de la population du Québec Les statistiques*: Institut national de santé publique Québec; 2006.

266. Davis LE, Mahar AL, Strumpf EC. Agreement between individual and neighborhood income measures in patients with colorectal cancer in Canada. *J Natl Cancer Inst*. 2023;115(5):514-22. DOI: https://doi.org/10.1093/jnci/djad017

267. Spiers GF, Liddle JE, Stow D, et al. Measuring older people's socioeconomic position: a scoping review of studies of self-rated health, health service and social care use. *J Epidemiol Community Health*. 2022;76(6):572-9. DOI: 10.1136/jech-2021-218265

268. Hu C-Y, Xing Y, Cormier JN, Chang GJ. Assessing the utility of cancer-registry–processed cause of death in calculating cancer-specific survival. *Cancer*. 2013;119(10):1900-7. DOI: https://doi.org/10.1002/cncr.27968

269. Wu Y, Ai Z, Xu G. Marital status and survival in patients with non-small cell lung cancer: an analysis of 70006 patients in the SEER database. *Oncotarget*. 2017;8(61):103518-34. DOI: 10.18632/oncotarget.21568

270. Chen Y, Yao L, Chen Q, et al. A retrospective study on the impact of radiotherapy on the survival outcomes of small cell lung cancer patients based on the SEER database. *Sci Rep*. 2024;14(1):15552. DOI: 10.1038/s41598-024-65314-8

271. David EA, Cooke DT, Chen Y, Nijar K, Canter RJ, Cress RD. Does Lymph Node Count Influence Survival in Surgically Resected Non-Small Cell Lung Cancer? *The Annals of Thoracic Surgery*. 2017;103(1):226-35. DOI: https://doi.org/10.1016/j.athoracsur.2016.05.018

272. Ryerson AB, Eheman C, Styles T, Rycroft R, Snyder C. Connecting the Dots: Linking the National Program of Cancer Registries and the Needs of Survivors and Clinicians. *Am J Prev Med*. 2015;49(6, Supplement 5):S528-S35. DOI: https://doi.org/10.1016/j.amepre.2015.08.026

273. Vaccarella S, Georges D, Bray F, et al. Socioeconomic inequalities in cancer mortality between and within countries in Europe: a population-based study. *Lancet Reg Health Eur*. 2023;25:100551. DOI: 10.1016/j.lanepe.2022.100551

274. Clegg LX, Reichman ME, Miller BA, 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(4):417-35. DOI: 10.1007/s10552-008-9256-0

275. Smailyte G, Jasilionis D, Ambrozaitiene D, Stankuniene V. Educational inequalities in cancer incidence and mortality in Lithuania: A record linkage study. *Cancer Epidemiol*. 2012;36(5):e279-e83. DOI: https://doi.org/10.1016/j.canep.2012.05.009

276. Teng AM, Atkinson J, Disney G, et al. Ethnic inequalities in cancer incidence and mortality: census-linked cohort studies with 87 million years of person-time follow-up. *BMC Cancer*. 2016;16(1):755. DOI: 10.1186/s12885-016-2781-4

277. Loehrer AP, Green SR, Winkfield KM. Inequity in Cancer and Cancer Care Delivery in the United States. *Hematol Oncol Clin North Am*. 2024;38(1):1-12. DOI: 10.1016/j.hoc.2023.08.001

278. OECD. *Beating Cancer Inequalities in the EU*2024.

279. Canadian Partnership Against Cancer. Pan-Canadian Cancer Data Strategy. 2023.

280. Persaud N, Sabir A, Woods H, et al. Preventive care recommendations to promote health equity. *Can Med Assoc J*. 2023;195(37):E1250. DOI: 10.1503/cmaj.230237

281. Kopp A, Kumar MB, Aitken N. COVID-19 mortality among First Nations people and Métis in private dwellings in Canada: An analysis of social determinants of health and health inequalities. Statistics Canada; 2024.

282. Malagón T, Yong JHE, Tope P, Miller WH, Jr., Franco EL. Predicted long-term impact of COVID-19 pandemic-related care delays on cancer mortality in Canada. *Int J Cancer*. 2022;150(8):1244-54. DOI: 10.1002/ijc.33884

283. Negoita S, Chen H-S, Sanchez PV, et al. Annual Report to the Nation on the Status of Cancer, part 2: Early assessment of the COVID-19 pandemic's impact on cancer diagnosis. *Cancer*. 2024;130(1):117-27. DOI: https://doi.org/10.1002/cncr.35026

284. Coronado AC, Finley C, Badovinac K, Han J, Niu J, Rahal R. Discrepancies Between Canadian Cancer Research Funding and Site-Specific Cancer Burden: A Spotlight on Ten Disease Sites. *Current Oncology*. 2018;25(5):338-41. DOI: https://www.mdpi.com/1718-7729/25/5/4230

285. Branswell B. An understudied killer: New network tackles lung cancer. McGill News. 2018.