Title: Quantifying the contribution of modifiable risk factors to socioeconomic inequities in cancer morbidity and mortality: a nationally representative population-based cohort study

Authors: Sara Nejatinamini[1], Jenny Godley[2], Leia M Minaker[3], Tolulope T Sajobi[1], Gavin R McCormack[1], Martin J Cooke[3], Candace IJ Nykiforuk[4], Lawrence de Koning[5], Dana Lee Olstad[1*]

Affiliations: [1]Department of of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada, [2]Department of Sociology, University of Calgary, Calgary, Alberta, Canada [3]School of Planning, University of Waterloo, Waterloo, Ontario, Canada, [4]School of Public Health, University of Alberta, Edmonton, Alberta, Canada, [5]Department of Pathology and Laboratory Medicine, University of Calgary, Calgary, Alberta, Canada.

*__Corresponding author:__ Dr Dana Olstad, Department of of Community Health Sciences, Cumming School of Medicine, University of Calgary, 3E16, 3280 Hospital Drive NW Calgary, AB, Canada, T2N 4Z6. Email: dana.olstad@ucalgary.ca; Phone: 403-210-8673.

Abstract word count: 338

Paper word count: 4800

Abstract

Background: Compared to those with a higher socioeconomic position (SEP), individuals with a lower SEP have higher cancer morbidity and mortality. However, the contribution of modifiable risk factors to these inequities is not known. This study aimed to quantify the mediating effects of modifiable risk factors to associations between SEP and cancer morbidity and mortality.

Methods: This study used a prospective observational cohort design. We combined eight cycles of the Canadian Community Health Survey (2000/2001–2011) as baseline data to identify a cohort of adults (≥35 years) without cancer at the time of survey administration (n =309,800). The cohort was linked to the Discharge Abstract Database (DAD) and the Canadian Mortality Database (CMDB) for cancer morbidity and mortality ascertainment. Individuals were followed from the date they completed the CCHS until 31 March 2013. Dates of individual first hospitalizations for cancer and deaths due to cancer were captured during this time period. SEP was operationalized using a latent variable combining measures of education and household income. Self-reported modifiable risk factors, including smoking, excess alcohol consumption, low fruit and vegetable intake, physical inactivity, and obesity were considered as potential mediators. Generalized structural equation modeling was used to estimate the mediating effects of modifiable risk factors in associations between low SEP and cancer morbidity and mortality in the total population and stratified by sex.

Results: Modifiable risk factors together explained 45.6% of associations between low SEP and overall cancer morbidity and mortality. Smoking was the most important mediator in the total population and for males, accounting for 15.5% and 40.2% of the total effect, respectively. For females, obesity was the most important mediator.

Conclusions: Modifiable risk factors are important mediators of socioeconomic inequities in cancer morbidity and mortality. Nevertheless, more than one half of the variance in these associations remained unexplained. Midstream interventions that target modifiable risk factors may help to alleviate inequities in cancer risk in the short-term. However, ultimately, upstream interventions that target structural

45 determinants of health are needed to reduce overall socioeconomic inequities in cancer morbidity and
46 mortality.

47 **Key words:** Socioeconomic position; modifiable risk factors; cancer morbidity and mortality; mediation
48 analyses

49 **Key messages**

50 • This was the first prospective, population-based study to quantify the mediating effects of
51 modifiable risk factors in associations between low socioeconomic position (SEP) and overall
52 cancer morbidity and mortality.
53 • Modifiable risk factors accounted for 45.6% of associations between low SEP and cancer
54 morbidity and mortality.
55 • Smoking was the most important mediator of associations between low SEP and cancer morbidity
56 and mortality both in the total population and for males, whereas obesity was the most important
57 mediator for females.
58 • While midstream interventions that target modifiable risk factors may help to alleviate inequities
59 in cancer morbidity and mortality in the short-term, ultimately upstream interventions that
60 address the inequitable distribution of power and resources within society are needed.

61 ## Introduction

62 According to the Canadian Cancer Society, nearly one in two Canadians will develop cancer in their
63 lifetime and one in four will die of the disease (1). Substantial progress has been achieved in cancer
64 control in Canada over the last few decades, including declines in the age-standardized incidence and
65 mortality rates of some of the most common cancers (1). Nevertheless, these advancements have not
66 benefitted all population groups equitably, as marked inequities in cancer morbidity and mortality have
67 been documented in Canada and other nations according to socioeconomic position (SEP) (2–5). The
68 World Health Organization's Commission on the Social Determinants of Health (CSDH) framework
69 positions SEP as a fundamental cause of health inequities because it shapes exposure to intermediary

determinants of health such as material circumstances, psychosocial factors, behavioral and biological risk factors, and access to the healthcare system, all of which collectively shape health (6).

This study is situated in the context of the CSDH framework to explore intermediary modifiable determinants of inequities in cancer morbidity and mortality, including smoking, excess alcohol consumption, low fruit and vegetable intake, physical inactivity, and obesity. These modifiable risk factors have long been recognized as important contributors to cancer. Indeed, it is estimated that approximately 24% of overall cancer incidence and 30% of overall cancer deaths worldwide are attributable to modifiable risk factors (7–9). However, very few studies have examined whether these factors mediate socioeconomic inequities in cancer incidence and mortality.

To our knowledge, just two studies have investigated whether modifiable risk factors might mediate associations between SEP and cancer morbidity and mortality. Hastert et al (10) examined the contribution of modifiable risk factors to inequities in cancer mortality among 54,737 older American adults (age 50–76). Modifiable risk factors, including body mass index (BMI), physical activity, diet quality, alcohol intake, smoking, and cancer screening accounted for 45% of the association between neighbourhood-level SEP and cancer mortality, with smoking explaining the greatest proportion of these associations (10). In another study, Doubeni et al (11) found that diet quality was a more important mediator of associations between SEP (operationalized using neighbourhood-level SEP and individual-level education) and risk of colorectal cancer in older American adults (age 50-71; n=506,488) compared to physical inactivity and smoking (11). In these studies, low response rates, non-representative sampling, and higher loss to follow up likely led to underrepresentation of low SEP groups, given that they are less likely to participate in research and more likely to drop out over the course of follow-up (12), potentially attenuating effect sizes. Nationally representative surveys linked to administrative health data may help to overcome these limitations by ensuring adequate representation of low SEP groups. The generalizability of these findings to other countries is also unclear as political, socioeconomic, and health care contexts differ between nations. Ideally, such studies should be conducted in a variety of nations to examine

pathways linking SEP with cancer in distinct national contexts. Such contextualized studies can provide important policy lessons for the global community.

In addition, these previous studies quantified mediation using a series of multiple regression equations, which do not allow for the simultaneous evaluation of predictors as mediators (10,11). Modifiable risk factors often cluster together, thus it is essential to use analytic methods such as structural equation modelling that can assess the simultaneous effects of multiple risk factors on cancer morbidity and mortality (13). In addition, combining multiple individual-level indicators of SEP such as income and education into a single latent variable within a structural equation model may better capture the complex construct of SEP than any single indicator in isolation (14,15).

The aim of this study was to examine (i) whether and to what extent modifiable risk factors, including smoking, excess alcohol consumption, low fruit and vegetable intake, physical inactivity, and obesity mediate associations between SEP and risk of overall cancer morbidity and mortality in a nationally representative sample of Canadian adults, and (ii) whether these relationships differed for males and females. For simplicity, we refer to all of these risk factors as modifiable risk factors throughout, while acknowledging that factors such as obesity are not necessarily readily or easily modifiable (16). We deliberately use the term modifiable as opposed to behavioral, first because obesity is not a behavior and second because we wish to highlight the social patterning of these factors rather than suggesting that they are a matter of individual choice.

Methods

Study design and cohort

This study used a population-based prospective observational cohort design whereby participants completed a cross-sectional survey and were subsequently followed longitudinally for cancer morbidity and mortality using administrative health/mortality data. The cohort consisted of adults who participated in the cross-sectional Canadian Community Health Survey (CCHS) at any point between 2000/2001-2011

and consented to data linkage (17) (Figure 1). Participants were included if they were at least 35 years of age and did not self-report cancer or pregnancy/breastfeeding at the time of survey administration. The follow-up period for each participant extended from the completion date of the cycle of the CCHS to which the participant responded (baseline) until the date of the earliest of the following events: first cancer hospitalization, death, or the endpoint of the study (31 March 2013).

Data sources

We used a dataset that linked respondents from eight cycles of the nationally representative, cross-sectional CCHS (2000/2001 - 2011) to longitudinal administrative health/mortality data in the Discharge Abstract Database (DAD; 2000-2013) and the Canadian Mortality Database (CMDB; 2000-2013) (18). Common identifiers were used to link consenting CCHS respondents (85.3%) to their administrative health data in the DAD (date of birth, sex, postal code, province issuing health information number and health information number of patients) and the CMDB (names, date of birth, sex and postal code). In the CCHS linked data, specific sampling weights were created to adjust for those who did not consent to share and link their data. Internal and external validations confirmed accuracy of the linkage process (19,20). Individuals may be represented more than once within and across the CCHS cycles. To avoid double counting, respondents with multiple records were identified and flagged in order to link them to the same death and hospitalization record (21).

Canadian Community Health Survey

The CCHS is a series of nationally and provincially representative cross-sectional surveys that use a multistage, stratified cluster sampling strategy to collect health-related information from individuals, including health status, health care utilization, and health determinants (17). The CCHS began collecting data in 2000/2001 and was repeated every two years until 2007, after which data were collected annually. While samples of approximately 130,000 respondents were interviewed in 2000/2001, 2003 and 2005, the sample size was changed to 65,000 respondents each year starting in 2007. Details about the CCHS

methodology have been described in detail elsewhere (17). Briefly, the CCHS collects data from Canadians aged 12 or older residing in a dwelling in the ten Canadian provinces and three territories and represents approximately 98% of the population. Individuals living on First Nations reserves or Crown land, in institutions, in remote regions, or who are full-time members of the Canadian Armed Forces are not included in the survey. The response rate across cycles ranged from 69.8% to 84.7%.

Discharge Abstract Database

The DAD captures administrative and clinical data for all patients discharged from acute care hospitals in Canada excluding respondents in Quebec and respondents from Manitoba before 2004, as Manitoba started reporting to the DAD as of April 1, 2004 (22). These patients were therefore excluded from the cohort (Figure 1). For each record, one "most responsible diagnosis" and up to 24 "secondary diagnoses" for hospital admission are coded according to the International Classification of Disease/Canadian Classification of Health Interventions (ICD/CCI) framework (22). For this study, admission date and diagnosis code (25 occurrences) was extracted for each hospitalization for each participant for all hospitalizations in the follow-up period. Overall cancer diagnoses are coded as malignant neoplasms 140-209 in ICD-9 and C00-D48 in ICD-10-CA. For consistency across survey cycles, ICD-9 codes were converted to ICD-10-CA/CCI codes where necessary (22).

Canadian Mortality Database

The CMDB collects cause of death information annually from all provincial and territorial vital statistics registries in Canada. For this study, date of death and cause of death were extracted. Death due to cancer was consistently coded using ICD-10CA as C00-D48.

Cohort creation

Linked CCHS, DAD and CMDB data files were merged in a two-step process. First, eight CCHS cycles corresponding to survey years 2000/2001, 2003, 2005, and 2007–2011 were combined and treated as baseline data (sample size of n=614,800 prior to exclusions). These pooled data were merged with CCHS-

DAD and CCHS-CMDB merge keys using household and person identification variables. These data were then merged with DAD hospitalization records and CMDB death records from 2001–2012/2013 using the same household and person identification variables, to create a dataset of respondent records containing CCHS, DAD and CMDB variables.

Data collection

Exposures

The exposure of interest was SEP which was operationalized as a latent variable, derived by combining annual household income and individual educational attainment. Participants in the CCHS reported total gross household income from all sources during the past 12 months. For each respondent, the adjusted ratio of their total household income to the Low Income Cut-Off corresponding to their household and community size was used to derive the distribution of household income and divided into deciles. Household income deciles were subsequently divided into two categories: low and lower-middle (deciles 1–5) and upper-middle and high (deciles 6–10). The highest level of education attained by respondents was dichotomized as post-secondary or less and greater than post-secondary education.

Mediators

Potential mediators in our analyses were smoking, excess alcohol consumption, low fruit and vegetable intake, physical inactivity, and obesity captured by the CCHS (10,11).

Total fruit and vegetable intake was based on the usual number of times per day that respondents reported consuming fruit, green salad, potatoes, carrots, and other vegetables, excluding juice, french fries, fried potatoes, or potato chips (23). The fruit and vegetable module in the CCHS is based on an existing validated module used in the National Cancer Institute's Dietary Screener Questionnaire (24). Low fruit and vegetable consumption was defined as less than 5 servings (400 g) of fruits and vegetables per day (25). A study by Traynor et al (2006) showed that self-reported number of times fruits and

vegetables were consumed daily was associated with the number of servings of fruits and vegetables consumed per day, and therefore we assumed that each time fruit or vegetable consumption was reported it was equivalent to one serving (23).

Respondents reported participation in leisure-time physical activity over the past 3 months. Self-reported participation in moderate-to-vigorous leisure-time physical activity is moderately correlated with measured physical activity among Canadian adults (26). A physical activity index was calculated by multiplying the number of times each activity was performed by the average duration of the activity by its energy cost (kilocalories/kg/hour) (17). Participants were classified as active/moderately active ($\geq$ 1.5 kcal/kg/day), or inactive (<1.5 kcal/kg/day). Alcohol intake was dichotomized as meeting or not meeting weekly Canadian recommendations (Women: < 10 drinks/week; Men: < 15 drinks/week) (27). Self-reported tobacco smoking was dichotomized as current/former smoker versus never (28,29). Self-reported smoking status has been shown to provide a valid estimate of the prevalence of smoking in Canada (30). Obesity was assessed based on BMI calculated from self-reported body weight and height. Obesity was defined as a BMI of $\geq$ 30 kg/m^2.

Outcome

The primary outcome was cancer morbidity and mortality (all types of cancer), derived by combining morbidity and mortality data and dichotomized as presence or absence of cancer morbidity and/or mortality (31–33). Cancer morbidity was defined from the DAD as presence or absence of hospitalization for cancer (most responsible diagnosis and all secondary diagnoses) during the follow-up period. Deaths due to cancer were identified through the CMDB.

Potential confounders

Potential confounders included in statistical models were sex (male, female), age (continuous variable) and race/ethnicity (White, Black, Chinese/Korean/Japanese, South /Southeast Asian, West Asian/Arab, Latin American, Multiple origins, Aboriginal only, other) (10).

Data analysis

Statistical analysis

Descriptive statistics were analyzed for the sex-stratified groups. A generalized structural equation modeling (GSEM) approach to path analysis was used to estimate mediating effects and to simultaneously test associations between low SEP, modifiable risk factors and cancer morbidity and mortality (13). As depicted in Figure 2, GSEM was used to test individual indirect effects via associations between SEP and potential mediators (a_1– a_4), and associations between potential mediators and cancer morbidity/mortality (b_1– b_4), controlling for SEP and potential confounders. The total indirect effect of SEP on cancer morbidity and mortality was assessed, where a*b represents the total individual indirect effect of SEP on cancer morbidity and mortality via each potential mediator. The direct effect, denoted by c', is the effect of SEP on cancer morbidity and mortality independent of all potential mediators and confounders. The proportion mediated by each risk factor was calculated as the risk factor-specific indirect effect divided by the total SEP-cancer effect ($a*b/(a_{1–4} * b_{1–4} + c)$, where the total effect represents the sum of the total indirect (indirect effect of all mediators combined) and direct (unmediated) effects.

Three separate pathway models were conducted to evaluate mediating effects of modifiable risk factors in associations between SEP and cancer morbidity and mortality in the total population (Model A), females (Model B), and males (Model C). Furthermore, to investigate potential exposure-mediator interaction, we added an exposure-mediator interaction term to our models (34). Models were fitted using the maximum likelihood method assuming logit links and Bernoulli distributions for all variables. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were used in model

comparisons (model with categorical mediators compared to model with binary mediators), with smaller values indicating a better fitting model.

As the follow-up time could vary for each cohort member, in a sensitivity analysis we conducted time-to-event analysis to take into account differing observation times. In this GSEM model, a Weibull distribution with a log link was used to handle time to cancer morbidity and mortality. Exponentiated coefficients were interpreted as hazard ratios.

Sex, age and race/ethnicity were included as potential confounders in all statistical models. The *svyset* command in Stata and Statistics Canada's bootstrap weights were used to generate unbiased estimates with variances adjusted for the sampling method (35). Missing values related to household income were imputed by Statistics Canada for all CCHS annual cycles except Cycle 1.1 (17). The percentage of missing values for other variables in our dataset were as follows: education 1.2%, income 4.0%; BMI 6.8%; fruit and vegetable intake 7.9%; physical activity 3.4%; alcohol intake 15%, race/ethnicity 1.0%. GSEM handles missing values using full information maximum likelihood estimation (36). All analyses were conducted with Stata version 15 (Stata Corp, Texas, USA), and statistical significance was assessed at $p < 0.05$.

Results

A total of 309,800 individuals met all study inclusion criteria and were included in the analyses (Figure 1). The mean age of the population was 54.0 (SD: 13.3) years. Nearly 36 % had less than a post-secondary level of education, while 44.2% of participants were classified as low/lower-middle income. Most (66.9%) reported consuming fruits and vegetables < 5 times daily. The percentage of participants with obesity was 17.9% and nearly one half of participants were physically inactive. One-third of participants reported never smoking, while about 80% reported meeting Canadian recommendations for alcohol intake.

(Table 1 here)

During a mean follow-up time of 7.2 ± 2.3 years, nearly 5.3% of participants were discharged from hospital with a cancer diagnosis and/or died from cancer (n = 21,565). Overall, cancer was responsible for 30.3% of deaths and 11.2% of the most responsible diagnoses for hospitalization. Cancer morbidity and mortality was higher among males than among females (5.6% *vs.* 5.0%; p<0.001).

The interaction coefficients for exposure-mediator interaction were non-significant. Models with binary variables had lower AIC/BIC values compared to models with categorical variables and therefore only binary variables were included in the reported models (AIC/BIC: 2441138 / 2441500 vs 2712874 / 2713257).

Direct and indirect effects

Model A was estimated in the total population and indicated a direct effect of low SEP on cancer morbidity and mortality (β= 0.11; 95% CI: 0.04-0.13), accounting for 54.4% of the total effect of low SEP. Overall, modifiable risk factors explained 45.6% of associations between low SEP and cancer morbidity and mortality. Smoking mediated 15.5% of associations between low SEP and cancer morbidity and mortality (p<0.001) (Table 2, Figure 2A).

Model B was estimated in females and indicated a direct effect of low SEP on cancer morbidity and mortality (β= 0.13; 95% CI: 0.01-0.14). Obesity was the most important mediator of associations between low SEP and cancer morbidity and mortality in females (p<0.001). The proportion mediated could not be estimated due to inconsistent mediation (Table 2, Figure 2B).

Model C was estimated in males and indicated a direct effect of low SEP on cancer morbidity and mortality (β= 0.11; 95% CI: 0.04-0.17), accounting for 45.2% of the total effect of low SEP on cancer morbidity and mortality. Current/former smoking was the most important mediator of associations between low SEP and cancer morbidity and mortality in males (40.2% of total effect mediated, p<0.001; Table 2, Figure 2C).

(Table 2 here)

Odds ratios

The likelihood of being discharged from hospital with a cancer diagnosis and/or dying from cancer was higher among respondents with a lower SEP and current/former smokers as compared to those with a higher SEP and those who never smoked (Table 3). Furthermore, females with obesity exhibited 1.13-fold higher odds for cancer morbidity and mortality (p=0.008).

(Table 3 here)

Sensitivity analysis

Results from the analysis with time-to-event as outcome were similar to results of the main analyses with respect to patterns of mediation. In addition, the hazard ratios were similar to the odds ratios from the main analysis. The hazard ratio of being discharged from hospital with a cancer diagnosis and/or dying from cancer was higher among respondents of low SEP and current/former smokers as compared with those of higher SEP and those who never smoked (Table 4). Obesity was associated with a higher hazard ratio of cancer morbidity and mortality in females (HR=1.21; CI=1.09-1.35) and in the total population (HR=1.10; CI=1.02-1.19).

(Table 4 here)

Discussion

We quantified the mediating role of modifiable risk factors in associations between low SEP and morbidity and mortality from all cancers combined using a nationally representative dataset containing linked survey and administrative data for 309,800 Canadian adults (≥35 years) in the context of the CSDH framework. Our results showed socioeconomic inequities in overall cancer morbidity and mortality in the Canadian population. Modifiable risk factors mediated 45.6% of associations between low SEP and cancer morbidity and mortality. Smoking was the most important mediator in both the total population and in males, mediating 15.5% and 40.2% of inequities in overall cancer morbidity and mortality, respectively. A different pattern was observed among females whereby obesity was the most important mediator of associations between low SEP and cancer morbidity and mortality.

Our results contribute to the evidence indicating socioeconomic inequities in cancer morbidity and mortality in developed nations and point to pathways underlying these associations (37–39). This is the first study to examine the mediating effects of modifiable risk factors in associations between a latent indicator of SEP with cancer morbidity and mortality using nationally representative data. We found that modifiable risk factors together accounted for 45.6% of these associations. In a previous study, Doubeni et al. found that modifiable risk factors (physical inactivity, unhealthy diet, smoking, and unhealthy weight) explained 43.9% of associations between education and risk of colorectal cancer in the United States (11). Hastert et al (2016) reported that modifiable risk factors (BMI, physical activity, diet quality, alcohol intake, smoking and cancer screening) explained 45% of associations between area-level SEP and cancer mortality in the United States (10). Overall, our findings are therefore consistent with prior studies that have examined the extent to which modifiable risk factors mediate associations between SEP with cancer morbidity and mortality separately.

Smoking mediated a considerable proportion of the observed associations between SEP and cancer morbidity and mortality in the total population and in males. These findings are substantiated by other studies (10,40). For instance, Hastert et al (2016) found that smoking mediated the greatest proportion (29%) of associations between area-level SEP and cancer mortality, compared to other modifiable risk factors including diet, physical activity, cancer screening and BMI. In the Whitehall II cohort study, Stringhini et al (2010) reported that smoking mediated 32% of associations between SEP (occupation) and total mortality; however, they did not find any associations between SEP and mortality due to cancer (40). The importance of smoking as a mediator in associations between SEP and cancer likely relates to the fact that individuals with a lower SEP are more likely to smoke compared to those with a higher SEP (41), and that smoking is a primary risk factor for multiple cancers (8).

Mediation models in females revealed evidence of inconsistent mediation, whereby the direct (c') and indirect effects (a*b) were opposite in sign and mediators acted as suppressor variable, nullifying the total effect (42). For instance, the indirect effect for smoking was negative because low SEP was associated with lower levels of smoking in females, while the direct effect was positive, and this

decreased the total effect for smoking. These differences in the mediating role of smoking caused by sex/gender-specific associations between SEP and smoking, with lower SEP associated with a higher prevalence of smoking in males, but with a lower prevalence among females are consistent with previous studies (41,43,44). Excess alcohol intake also acted as a suppressor mediator in females, as associations between low SEP and excess alcohol intake had a negative sign, and thus an opposite effect compared to the direct effect, which was positive, nullifying the total effect for alcohol. Similarly, findings from national survey data from 13 countries including Canada showed that lower SEP women were less likely than higher SEP women to drink excessively, while the opposite was observed in men in most countries (45).

Obesity was the most important mediator of SEP-related cancer morbidity and mortality in females, whereas it did not mediate these associations in males. To our knowledge, no prior research has examined differences in obesity-mediated cancer inequities according to sex/gender. Although participants in the CCHS were only asked to report biological sex, biological sex and socially constructed gender roles are inextricably linked, and thus it was not possible to disentangle their effects in this study. Previous research in Canada and the US supports the existence of socioeconomic and sex/gender inequities in obesity (46,47). Obesity is more frequently observed among females with a lower SEP, whereas in males obesity is more concentrated among those with a higher SEP (particularly higher income) (48–51). Sex/gender differences in the socioeconomic patterning of obesity may arise from cultural and symbolic values of body shape whereby larger female body shapes are stigmatized, and the larger body sizes of males are considered to indicate strength, power and dominance (48,52). Thus, our results are consistent with these sexed/gendered socioeconomic differentials in obesity, and the known contribution of obesity to risk of multiple cancers (8).

Given that modifiable risk factors mediated nearly one half of associations between low SEP and cancer morbidity and mortality in the total population, our results suggest that low SEP may confer vulnerability to cancer in part by diminishing access to economic, cultural and social resources that can

promote and support positive lifestyle practices related to smoking, alcohol intake, physical activity, dietary intake, and body weight (53). Midstream interventions that address these modifiable risk factors (e.g. targeted healthy food subsidies, locating recreation and sports facilities in disadvantaged neighbourhoods) may therefore represent one important opportunity to reduce inequities in cancer in the short-term. Such interventions are particularly important given the difficulty and long timelines required to implement more upstream interventions. Nevertheless, more than one half of the total effect of SEP remained unexplained and there was some inconsistent mediation in females. In this respect, our results challenge conventional thinking that inequities in cancer risk are simply the result of "risky lifestyle behaviours" on the part of socioeconomically disadvantaged groups and indicate the importance of examining other potential mediators. Future studies might assess the mediating role of social support, social capital, resilience, discrimination, housing, stress, and/or food insecurity, to name a few. For instance, food insecurity was the strongest mediator of associations between county-level median incomes and cancer mortality in one US study (54).

It is important to acknowledge that in referring to dietary intake, obesity, excess alcohol consumption, smoking and physical inactivity as modifiable risk factors, we do not mean to imply that they result from the poor choices of socioeconomically disadvantaged groups. Systematic differences in cancer morbidity and mortality between low and high SEP groups are a consequence of social contexts that shape the distribution of resources within society according to factors such as level of income, educational attainment, occupation and others (55,56). That is, people adopt particular lifestyle patterns according to their assessments of what is structurally possible for them, such as with respect to the costs of fruits and vegetables in relation to available household income (57). Thus, the social context both empowers and constrains the exercise of human agency, and it is only by addressing contextual constraints that individuals can exercise their agency in health-promoting ways (58). Therefore, the extent to which these factors are modifiable by individuals is limited. Rather, these factors are more appropriately portrayed as modifiable via structural change. As discussed in the CSDH framework, while midstream interventions that address intermediary modifiable determinants among lower SEP groups may

help to alleviate inequities in cancer risk in the short-term, ultimately upstream structural interventions that address the inequitable distribution of power and resources within society are needed to substantially reduce socioeconomic inequities in cancer morbidity and mortality. Such strategies might encompass policies that ensure universal access to high quality childcare and education, legislating a living wage, increases to social assistance rates; or more radically a universal basic income guarantee.

This study demonstrates several important strengths, including its large sample size, use of nationally representative data, and prospective design with objectively measured, long-term health outcomes for all participants. Excluding participants with a history of cancer at baseline helped to minimize reverse causality. Measurement of mediating effects using GSEM is another strength, as GSEM enables simultaneous consideration of multiple potential mediators and pathways and has much greater statistical power than standard regression models (13). Limitations include the potential for social desirability, non-response and/or recall biases due to the self-reported nature of exposures and mediators in the survey data. However, survey weights were applied to account for non-response bias. As respondents' sociodemographic and risk factor information was collected at baseline only, it was not possible to measure changes in SEP and modifiable risk factors over the course of follow-up. We combined information on the two most common measures of SEP, education and household income, to create a latent SEP variable to capture a more comprehensive perspective of SEP. Nevertheless, SEP is a complex construct that is not adequately captured by these two, or any other combination of indicators (59). Therefore, future studies are needed to understand how other dimensions of SEP interact to shape risk of cancer morbidity and mortality. Although we had little missing data overall, 15% of data pertaining to alcohol intake were missing, which could bias these estimates. In addition, respondents from Quebec were excluded from the cohort, as Quebec does not report to the DAD, while residents of Manitoba were excluded prior to 2004. These exclusions could limit the representativeness of our results. There are different practices in the literature with respect to the use of time lags. While some studies opt to exclude the first year of follow-up, we followed the practices of others (11,40) who have included the

first year of follow-up given that our SEP exposures and modifiable risk factor mediators (e.g. smoking, obesity) were likely to have been stable for several years prior to their measurement at baseline. Moreover, the rate of cancer morbidity and mortality in the first year of follow-up was similar to the rate in all subsequent years. Finally, we focused on combined cancer morbidity and mortality as our outcome of interest; however, the total mediating effects of modifiable risk factors were comparable with two previous studies that examined these outcomes separately (10,11).

In conclusion, low SEP was associated with higher cancer morbidity and mortality in the Canadian population. Modifiable risk factors were important mediators of inequities in cancer morbidity and mortality, although more than one half of these associations remained unexplained. While midstream interventions that target modifiable risk factors may help to alleviate inequities in cancer risk in the short-term, ultimately upstream structural interventions are needed to prevent and reduce inequities in cancer morbidity and mortality.

Ethics approval

The Conjoint Health Research Ethics Board at the University of Calgary deemed this study exempt from ethical approval as it involved secondary data analysis of a survey conducted by Statistics Canada.

Funding

This study was funded by a Canadian Institutes of Health Research Operating Grant (FRN 156728) and a Petro-Canada Young Innovator Award in Community Health. SN was supported by a Libin Cardiovascular Institute/Cumming School of Medicine Postdoctoral Award.

Data availability

The data underlying this article were provided by Statistics Canada under licence / by permission. Data will be shared on request to the corresponding author with permission of Statistics Canada.

Conflict of interest

DLO has received research support from a Petro-Canada Young Innovator Award in Community Health. All other authors declare no conflict of interest.

Figure 1 Cohort selection
CCHS: Canadian Community Health Survey; DAD: Discharge Abstract Database

Figure 2. Causal pathway diagrams for evaluation of mediating effects of modifiable risk factors on associations between low socioeconomic position (SEP) and cancer morbidity and mortality in the total population (Model A), females (Model B), and males (Model C); The a_1–a_5, b_1–b_5 and c' represent path coefficients; Adjusted for sex, age, and race/ethnicity; *p<0.05; **p<0.001

References

1. Canadian Cancer Statistics 2017. Canadian Cancer Society, Toronto, Canada. https://www.cancer.ca/Canadian-CancerStatistics-2017-EN.pdf

2. Freeman HP. Poverty, Culture, and Social Injustice: Determinants of Cancer Disparities. CA Cancer J Clin 2004;54:72–7.

3. Ward E, Jemal A, Cokkinides V, et al. Cancer Disparities by Race/Ethnicity and Socioeconomic Status. CA Cancer J Clin 2004;54:78–93.

4. Willis K, Hajizadeh M. Socioeconomic inequalities in gastric cancer incidence in Canada: 1992–2010. Acta Oncol 2020;59:1333-1337.

5. Hajizadeh M, Johnston GM, Manos D. Socioeconomic inequalities in lung cancer incidence in Canada, 1992–2010: results from the Canadian Cancer Registry. Public Health 2020;185:189–95.

6. Solar O, Irwin A. A Conceptual Framework for Action on the Social Determinants of Health. Soc Determ Heal Discuss Pap 2 (Policy Pract) 2010;79.

7. Ott JJ, Ullrich A, Mascarenhas M, Stevens GA. Global cancer incidence and mortality caused by behavior and infection. J Public Health 2011;33:223–33.

8. Danaei G, Vander Hoorn S, Lopez AD, Murray CJ, Ezzati M. Causes of cancer in the world: Comparative risk assessment of nine behavioural and environmental risk factors. Lancet 2005;366:1784–93.

9. Parkin DM, Boyd L, Walker LC. The fraction of cancer attributable to lifestyle and environmental

461 factors in the UK in 2010. Br J Cancer 2011;105:S77–81.

462 10. Hastert TA, Ruterbusch JJ, Beresford SAA, Sheppard L, White E. Contribution of health
463 behaviors to the association between area-level socioeconomic status and cancer mortality. Soc Sci
464 Med 2016;148:52–8.

465 11. Doubeni CA, Major JM, Laiyemo AO, et al. Contribution of behavioral risk factors and obesity to
466 socioeconomic differences in colorectal cancer incidence. J Natl Cancer Inst 2012;19:1353–1362.

467 12. Unger JM, Gralow JR, Albain KS, Ramsey SD, Hershman DL. Patient income level and cancer
468 clinical trial participation: A prospective survey study. JAMA Oncol 2016;2:137–9.

469 13. Kupek E. Beyond logistic regression: Structural equations modelling for binary variables and its
470 application to investigating unobserved confounders. BMC Med Res Methodol 2006;6:13.

471 14. Martinez SA, Beebe LA, Thompson DM, Wagener TL, Terrell DR, Campbell JE. A structural
472 equation modeling approach to understanding pathways that connect socioeconomic status and
473 smoking. PLoS One 2018; 1;13(2).

474 15. Marmot M. Closing the gap in a generation. Health Equity Through Action Soc Determ Heal.
475 Lancet 2008; 372:1661-1669.

476 16. Farias MM, Cuevas AM, Rodriguez F. Set-point theory and obesity. Metabolic Syndrome and
477 Related Disorders 2011;9:85–9.

478 17. Statistics Canada. Canadian Community Health Survey—Annual Component (CCHS) Ottawa,
479 ON: Statistics Canada 2018.
480 https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3226

481 18. Population health survey data linked to hospitalization, mortality, emergency department visit,
482 cancers, and tax file: user guide. Ottawa: Statistics Canada; 2019. [Unpublished internal
483 document. Available upon request.]

19. Sanmartin C, Decady Y, Trudeau R, et al. Linking the Canadian community health survey and the canadian mortality database: An enhanced data source for the study of mortality. Heal Reports 2016;27:10–8.

20. Rotermann M. Evaluation of the coverage of linked Canadian Community Health Survey and hospital inpatient records. Health Rep 2009;20:45–51.

21. Sanmartin C, Decady Y, Trudeau R, Dasylva A, Tjepkema M, Finès P, et al. Linking the Canadian community health survey and the canadian mortality database: An enhanced data source for the study of mortality. Heal Reports 2016;27:10–8.

22. Canadian Institute for Health Information. Discharge Abstract Database metadata (DAD). Available from: https://www.cihi.ca/en/discharge-abstract-database-metadata

23. Traynor MM, Holowaty PH, Reid DJ, Gray-Donald K. Vegetable and fruit food frequency questionnaire serves as a proxy for quantified intake. Can J Public Heal. 2006;97:286–90.

24. National Cancer Institute (NCI). Dietary screener questionnaire in the NHANES 2009–10. http://epi.grants.cancer.gov/nhanes/dietscreen/. Updated October 15, 2015. Accessed July 12, 2016.

25. Norat T, Aune D, Chan D, Romaguera D. Fruits and vegetables: Updating the epidemiologic evidence for the WCRF/AICR lifestyle recommendations for cancer prevention. Cancer Treat Res 2014;159:35–50.

26. Colley RC, Butler G, Garriguet D, Prince SA, Roberts KC. Comparison of self-reported and accelerometer-measured physical activity in Canadian adults. Heal Reports 2018;29:3–15.

27. Canadian Centre on Substance Abuse. Canada's low-risk alcohol drinking guidelines [Internet]. Canadian Centre on Substance Abuse 2011:1–2. Available from: www.ccsa.ca

28. Zha L, Sobue T, Kitamura T, et al. Changes in smoking status and mortality from all causes and lung cancer: A longitudinal analysis of a population-based study in Japan. J Epidemiol. 2019;29:11–7.

29. Dias M, Linhas R, Campainha S, Conde S, Barroso A. Lung cancer in never-smokers–what are the differences? Acta Oncol 2017;56:931–5.

30. Wong SL, Shields M, Leatherdale S, Malaison E, Hammond D. Assessment of validity of self-reported smoking status. Health Rep 2012;23:47-53.

31. Méjean C, Droomers M, van der Schouw YT, et al. The contribution of diet and lifestyle to socioeconomic inequalities in cardiovascular morbidity and mortality. Int J Cardiol 2013;168:5190-5.

32. Marmot MG, Shipley MJ, Hemingway H, Head J, Brunner EJ. Biological and behavioural explanations of social inequalities in coronary heart disease: The Whitehall II study. Diabetologia 2008;51:1980–8.

33. Kim JL, Cho KH, Park EC, Cho WH. A single measure of cancer burden combining incidence with mortality rates for worldwide application. Asian Pacific J Cancer Prev 2014;15:433–9.

34. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods 2013;18:137–50.

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.

36. Cassim R, Milanzi E, Koplin JJ, Dharmage SC, Russell MA. Physical activity and asthma: Cause or consequence? A bidirectional longitudinal analysis. J Epidemiol Community Health

2018;72:770–5.

37. Hoebel J, Kroll LE, Fiebig J, et al. Socioeconomic Inequalities in Total and Site-Specific Cancer Incidence in Germany: A Population-Based Registry Study. Front Oncol 2018;8.

38. Mao Y, Hu J, Ugnat A-M, Semenciw R, Fincham S. Socioeconomic status and lung cancer risk in Canada. Int J Epidemiol 2001;30:809–17.

39. Vinnakota S, Lam NSN. Socioeconomic inequality of cancer mortality in the United States: A spatial data mining approach. Int J Health Geogr 2006;5.

40. Stringhini S, Sabia S, Shipley M, et al. Association of socioeconomic position with health behaviors and mortality. JAMA - J Am Med Assoc 2010;303:1159–66.

41. Corsi DJ, Lear SA, Chow CK, Subramanian S V., Boyle MH, Teo KK. Socioeconomic and Geographic Patterning of Smoking Behaviour in Canada: A Cross-Sectional Multilevel Analysis. PLoS One 2013;8:57646.

42. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation Analysis. Annu Rev Psychol 2007;58:593–614.

43. Bobak M, Jarvis MJ, Skodova Z, Marmot M. Smoke intake among smokers is higher in lower socioeconomic groups. Tob Control 2000;9:310–2.

44. Nicolaou SA, Heraclides A, Markides KS, Charalambous A. Prevalence and social determinants of smoking in the adult greek cypriot population. Hippokratia 2016;20:284–91.

45. Devaux M, Sassi F. Social disparities in hazardous alcohol use: self-report bias may lead to incorrect estimates. Eur J Public Health 2016;26:129-34.

46. Twells LK, Gregory DM, Reddigan J, Midodzi WK. Current and predicted prevalence of obesity in Canada: a trend analysis. C Open 2014;2:18–26.

552 47. Booth HP, Charlton J, Gulliford MC. Socioeconomic inequality in morbid obesity with body mass
553 index more than 40 kg/m2 in the United States and England. SSM - Popul Heal 2017;3:172–8.

554 48. McLaren L. Socioeconomic status and obesity. Epidemiologic Reviews 2007;29:29-48.

555 49. Gebreab SY, Diez-Roux A V., Hickson DMA, et al. The contribution of stress to the social
556 patterning of clinical and subclinical CVD risk factors in African Americans: The Jackson Heart
557 Study. Soc Sci Med 2012;75:1697–707.

558 50. Shields M, Tjepkema M. Trends in adult obesity. Health Rep 2006;17:53-9.

559 51. Godley J, Mclaren L. Socioeconomic Status and Body Mass Index in Canada: Exploring Measures
560 and Mechanisms. Can Rev Sociol 2010;47:381–403.

561 52. Fikkan JL, Rothblum ED. Is Fat a Feminist Issue? Exploring the Gendered Nature of Weight Bias.
562 Sex Roles 2012; 66:575–92.

563 53. Hiatt RA, Breen N. The Social Determinants of Cancer. A Challenge for Transdisciplinary
564 Science. American Journal of Preventive Medicine 2008; 35:141-50.

565 54. O'Connor JM, Sedghi T, Dhodapkar M, Kane MJ, Gross CP. Factors Associated With Cancer
566 Disparities Among Low-, Medium-, and High-Income US Counties. JAMA Netw open
567 2018;1:183146.

568 55. Closing the gap in a generation: health equity through action on the social determinants of health.
569 Final Report of the Commission on Social Determinants of Health. Geneva, World Health
570 Organization. CSDH (2008).

571 56. Phelan JC, Link BG, Tehranifar P. Social Conditions as Fundamental Causes of Health
572 Inequalities: Theory, Evidence, and Policy Implications. J Health Soc Behav 2010;5:28–40.

573 57. Solar, O. & Irwin A. A Conceptual Framework for Action on the Social Determinants of Health.

574 Discussion Paper for the Commission on Social Determinants of Health. World Health
575 Organisation. 2010. Available from:
576 http://www.who.int/sdhconference/resources/ConceptualframeworkforactiononSDH_eng.pdf

577 58. Marmot M, Allen JJ. Social determinants of health equity. American Journal of Public Health
578 2014;104:517-9.

579 59. Olstad DL, McIntyre L. Reconceptualising precision public health. BMJ Open. 2019;9:e030279.