McMaster University

STATure

STATistical Applications For Undergraduate Research Exploration

Volume 4 – Winter 2025

The field of statistics is the study of learning from data, but this course is much more than that. At the beginning of LifeSci 3LX3 (formerly 3LL3), I often hear that students feel nervous about the idea of working with data because they *aren't math people*. I hear that they feel anxious about the independence in this course, or how they didn't *get* other stats classes. Over the course of the semester, students start by learning about how to use SPSS, a popular statistics software program, and continue to weave in basic data analysis skills. They learn some theory, but the focus is on application within small groups. Some students feel uneasy about the group work in the course, as projects are completed with 1-3 other students, and with group work often comes challenges. But, many students have told me that this course helps them be better group members and gives them the chance to form friendships and create trusting relationships with their peers. I try to emphasize at the beginning of the course that the group work shouldn't be viewed as transactional, but I hope that they see the value in investing in the relationships that they form with their peers by communicating well and respectfully and most importantly, by showing up to class. So, these manuscripts in this issue represent a semester's worth of learning in statistics and scientific writing. But, these manuscripts are also a testament to student growth, resilience (this was an 8:30am class, after all), curiosity and being open to course that focuses more on application and skill building and less on memorizing PowerPoint slides! This **STATure** (**STAT**istical applications for **u**ndergraduate **r**esearch **e**xploration) issue is a record and celebration of the students' work from the Winter 2025 offering of LifeSci 3LX3.

Thank you to Saikeerthana Sundar, Zainab Amer and Julian Chan, students who were in LifeSci 3LX3, for creating cover art, collating and formatting all manuscripts for this issue. I also thank our supportive Teaching Assistant, Maryam Rahat, for helping me to coach our students in class and in many hours of extra help office hour sessions.

This is not a peer-reviewed journal. Students received feedback on their projects and writing throughout the semester, and manuscripts were submitted at the end of the semester. Some spelling and grammatical errors may appear.

I hope you enjoy reading these creative and innovative research projects that our students worked so hard on.

Dr. Janet Pritchard

Assistant Professor, School of Interdisciplinary Science

LifeSci 3LX3 course creator and instructor

Cover Art Acknowledgment

The cover art was created by Zainab Amer and is visually based on a stock illustration titled *Seamless background with math formulas and graphics on white*. The original image is available on iStock (https://media.istockphoto.com/id/586090030/vector/seamless-background-with-math-formulas-and-graphics-on-white.jpg) and is used here as visual inspiration under standard stock licensing terms.

Acknowledgment

Zainab Amer, Julian Chan, and Sai Sundar contributed to the final editing and compilation of the manuscript. Their attention to detail and commitment to producing a cohesive and polished document are sincerely appreciated.

Table of Contents

Better Learning, Safer Skin?

Exploring the Influence of Education Attainment on Sun Protection Behaviours

Djurdjina Gligoric, Yitao Liu, & Yovana Zubac

Faculty of Science, McMaster University

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8th, 2025

Introduction

Skin cancer is one of the most common forms of cancer in Canada. Rates of melanoma have been steadily increasing over the past several decades with over 80,000 cases diagnosed annually (Canadian Skin Cancer Foundation, 2018). In 2016, 6800 Canadians were diagnosed with melanoma, the deadliest form of skin cancer, and 1200 will die from it (Canadian Skin Cancer Foundation, 2018). In 2024, the number of cases diagnosed rose to 11,300 (Brenner et al., 2024). The primary risk factor for various types of skin cancer is prolonged exposure to ultraviolet (UV) radiation (Watson et al., 2024). This form of electromagnetic radiation is emitted by the sun with UV-B rays being particularly harmful to human health (Pinault & Fioletov, 2017). They cause direct DNA damage in skin cells, oxidative stress and immunosuppression, accelerating skin aging and collectively driving carcinogenesis (Narayanan et al., 2010). Everyday behaviours can influence our exposure to UV rays and important risk factors for skin cancer. For instance, those who experience several sunburns throughout their life, especially severe ones at a young age, have a higher lifetime risk of developing skin cancer (Amaro-Ortiz et al., 2014). Furthermore, geographic and environmental factors, including altitude, cloud cover, and ozone depletion, further influence individual exposure (Kalia et al., 2019). Additionally, climate change plays an increasingly important role by altering environmental conditions that affect UV intensity (Watson et al., 2024). Warmer weather can lead to more time spent outdoors, while changes in cloud patterns and pollution levels impact UV exposure across regions (Barnes et al., 2019). These evolving risks underscore the need for updated public health strategies that consider both behavioural and environmental factors influencing skin cancer prevention.

Although public awareness of UV dangers has grown, widespread adoption of protective behaviours remains inconsistent. A common misconception is that avoiding sun exposure could result in vitamin D deficiency; however, research indicates that nearly 40% of individuals with high sun exposure still experience deficiency (Kalia et al., 2019). Additionally, the effectiveness of protection strategies remains inconsistent across demographics. For instance, young adults aged 18–34, consistently report high rates of sunburn and low adherence to protective behaviours (Pinault & Fioletov, 2017). In one study, 94% of individuals aged 18–30 had experienced at least one sunburn in the past year (Lee et al., 2014). A report from the Centers for Disease Control and Prevention (CDC) highlighted that in adults aged 18–29, protective behaviours like seeking shade or using sunscreen are inconsistently practiced (CDC, 2012). Melanoma is now the second most common cancer among individuals aged 15–29, making them a crucial group to study with early-life sunburns significantly contributing to lifetime risk ("Facts & Statistics", 2025; Amaro-Ortiz et al., 2014). Conversely, individuals aged 55 and older face a heightened incidence of melanoma. Around half of melanoma cases occur in people over the age of 50, with 59 being the median age at diagnosis ("Age and Risk", 2024). Despite this increased risk, less than half of older adults consistently engage in sun protection behaviours when outdoors for extended periods (CDC, 2023). This discrepancy highlights the importance of understanding how age intersects with knowledge, perception, and behavioural responses to UV risk.

Another critical factor that may influence sun protection habits is education. Several studies suggest that higher levels of education are associated with increased awareness and adoption of sun-safe practices (Kalia et al., 2019; Falk & Anderson, 2013). However, greater knowledge does not always translate into consistent behaviour. For example, Almuqati et al. (2019) found that although nearly 79% of university students in Saudi Arabia recognized the harmful effects of excessive sun exposure, only 24% reported regularly using sunscreen. This disconnect between awareness and action raises questions about what drives preventive behaviours and whether educational level alone is a sufficient predictor.

Importantly, there is limited research on how education level influences sun protection behaviours in the Canadian context. Only one study, dating back to 1994 from Campbell and Birdsell found that while many Albertans acknowledged the risks of sun exposure, few took adequate precautions, and many believed occasional sunburns were harmless. Understanding the patterns of behaviour among demographics in Canada is crucial for the development of meaningful and useful public health strategies for mitigating exposure to UV rays, and subsequently, development of skin cancers. This study aims to explore the relationship between education level and self-reported sun protection behaviours to inform more targeted skin cancer prevention efforts in Canada. This leads us to ask; when comparing groups of participants with different education levels (did not complete secondary, completed secondary, or completed post-secondary) from a sample of Canadians 15 and over, is highest level of education associated with self-reported sun-protective precautions? As a follow up question, do self-reported sun-protective behaviours differ between younger adults (20-34) and older adults (55+) among the participants in the sample?

Methods

Study design

This cross-sectional study is a secondary analysis of data from the 1996 Canadian Sun Exposure Survey, conducted by Statistics Canada. The survey, retrieved from the Ontario Data Documentation, Extraction Service, and Infrastructure Initiative (ODESI), aimed to assess sun exposure attitudes and behaviours among Canadians aged 15 and older. Data was collected via a 55-item, random-digit-dialing telephone survey conducted between September and October 1996 at both the national and regional levels (Shoveller, et al., 2000). Survey questions asked individuals about their sun related behaviours during leisure, work time, and the winter season for the period of June to August 1996 (Purdue, 2002). While individuals aged 15 and older were the predominant focus for this survey, data was also collected on children under the age of 15, but this data was reported by their parents (Statistics Canada, 2023).

Study participants

The inclusion criteria included: (a) males and females, (b) individuals aged 15 and older, (c) Canadian citizens, and (d) individuals who are obtaining or obtained some level of education. The exclusion criteria for our study excluded cases who did not provide a yes or no response to the statement asking individuals if they take extra precautions when the UV rating is high (our dependent variable). We also excluded cases that answered, “do not know”, “refused” or “not stated” for our independent variable of highest level of education completed.

Study outcomes

The primary outcome for our study investigated individuals’ highest level of education obtained and their self-reported sun protection behaviours. This outcome was chosen as previous literature has highlighted the link between education and sun protection, but the correlation to actionable precautions still remains unclear (Falk & Anderson, 2013). The dependent variable was self-reported sun protection behaviour, defined by the survey question: “Do you take extra precaution when the UV rating is high?”, in which responses could be “Yes”, “No”, “Valid Skip”, “Do not know”, “Refused”, or “Not Stated”. Our primary independent variable was highest level of education attainment, in the sample of individuals 15 years of age and older, in which participants responded, “Has not completed secondary school”, “Completed secondary school”, “Completed post-secondary", “Do not know”, “Refused” or “Not Stated”.

For the secondary outcome, we narrowed down our sample to analyze whether self-reported sun protection behaviours between two defined age groups (younger adults vs. older adults) differed. This outcome was chosen since many studies highlighted that both younger and older adults have increasing risks of developing melanoma (“Age and Risk”, 2024), yet they fail to engage in effective sun protective behaviours (Pinault & Fioletov, 2017; CDC, 2023). The dependent variable remained the same as in the primary analysis. The independent variable for the secondary outcome was participants within the age categories of 20-34 or 55+. To obtain these age groups, our existing age categories needed to be transformed. The 20-24 and 25-34 groups were combined to produce the 20-34 category and the 55-64 and 65+ groups were combined to produce the 55+ category.

To refine the dataset, all missing responses, and responses that were of no interest to the study outcomes were removed. Responses to the dependent variable (Extra precaution when high UV rating) that were; valid skip, not stated, do not know, or refused, were removed from our dataset. Additionally, missing data for both independent variables were also removed from the dataset. For the primary question, all those who answered, “Do not know”, “Refused” or “Not Stated” for highest level of education attainment, were removed from the dataset. For our secondary question, cases that fit into the categories, “15-19”, “35-44”, and “45-54” were removed from the dataset. This left only cases within the categories “20-34” and “55+”.

Additional Data Collection

Other supplementary information was collected to obtain a better insight into the study population. These variables include sex (male and female), country in which individual was born (Canada or International), and household income over the last 12 months (No income to $29K, $30K-$69K, and $70K-$80K+). Previous literature has found that these 3 variables can contribute to individuals sun protective behaviours (Raymond-Lezman & Riskin, 2023).

Statistical analysis

All statistical analyses were conducted using IBM SPSS (Statistical Package for the Social Sciences) V.30. A normality test was not conducted for this study as no continuous variables were present in our dataset. Most variables in the dataset were categorical, therefore, descriptive statistics, including frequencies and percentages, were calculated for all variables and summarized in Table 1. A binary logistic regression was conducted to address our primary research question, as it allowed us to predict the probability of participants taking extra sun precautions (yes/no) based on their highest level of education categorization. For our secondary question, a chi-squared test was run as this enabled us to compare the between group differences in the proportion of responses to our categorical (yes/no) question based on age. Statistical significance was set at a p-value less than 0.05.

Results

Number of participants

Prior to removing missing cases our total sample size was 4023 participants. There were 489 missing cases for our dependent variable ("Do you take extra precautions when UV is high?"), which we deleted from our database for further analysis, leaving us with 3534 cases. We also excluded any cases that were missing for our independent variable of education level (n = 783 missing), which took us from 3534 cases to 2751 cases. Therefore, the total population sample size for our primary research question was 2751 participants. For our secondary question, we started off with a sample size of 2751 participants and after combining our 2 age groups together (20-24 and 25-34 and 55-64 and 65+), there were 1217 cases that were of no interest for our analysis. After deleting these 1217 missing cases, we ended up with a sample size of 1534 cases for our secondary analysis.

Descriptive Statistics

Descriptive statistics with basic demographic criteria for the 2751 participants are included in Table 1. Age group was transformed from 8 into 4 categories for clearer presentation. Household income was transformed from 10 into 3 categories, based on low, middle, and high annual household income in Canada, 1996 (Statistics Canada, 1998).

Table 1: Descriptive Characteristics of Sun Exposure Survey participants aged ≥15 years by the highest educational level completed (N = 2751).

	Did not complete secondary (n = 831)	**Completed secondary** (n = 1332)	**Completed post-secondary** (n = 588)
categorical variable presented as n (%)			
Age (years)			
15-19	131 (15.8)	91 (6.8)	8 (1.4)
20-34	120 (14.4)	425 (31.9)	261 (44.4)
35-54	217 (26.1)	527 (39.6)	243 (41.3)
55+	363 (43.7)	289 (21.7)	76 (12.9)
Sex			
Female	426 (51.3)	715 (53.7)	324 (55.1)
Male	405 (48.7)	617 (46.3)	264 (44.9)
Country in which they were born*			
Canada	761 (91.6)	1153 (86.7)	468 (79.6)
International	70 (8.4)	177 (13.3)	120 (20.4)
Household income in the last 12 ** **months**			
No income to $29K	418 (61.2)	453 (38.2)	151 (28.2)
$30K to $69K	231 (33.8)	589 (49.6)	292 (54.5)
$70K to $80K+	34 (5.0)	145 (12.2)	93 (17.3)

Data are presented as frequency (%) of participants. *Complete secondary n = 1330; ** Did not complete secondary n = 683; complete secondary n = 1187; complete post-secondary n = 536

Primary research question—bar graph and logistic regression

Figure 1 illustrates the distribution of individuals across different levels of education regarding their self-reported actions in response to high UV ratings, prior to any statistical analysis. Across all educational categories, more than half of the participants reported taking extra precautions when UV levels were high.

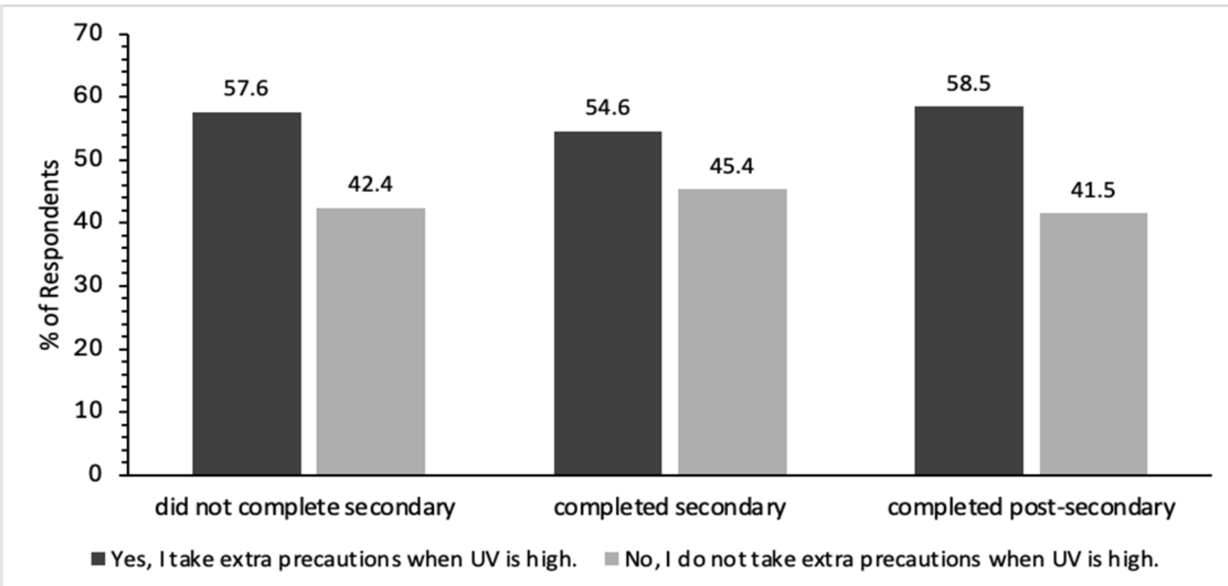


Figure 1. A double bar graph comparing the proportions of individuals who self-reported take or do not take extra precautions under high UV ratings based on their highest level of education attainment (N = 2751). Dark grey bars represent individuals who answered yes to the question "Do you take extra precautions when UV is high?" (n = 479 for did not complete secondary, n = 727 for completed secondary, n = 344 for completed post-secondary). Light grey bars indicate individuals who answered no to the same question (n = 352 for did not complete secondary, n = 605 for completed secondary, n = 244 for completed post-secondary).

Table 2 demonstrates the results of a binary logistic regression analysis examining the degree of association between highest education level achieved and the likelihood of taking extra sun precautions. Without the predictor variable (age group) in the model, the classification table for Block 0 showed 56.3% accuracy based solely on SPSS predictions. The Omnibus Tests of Model Coefficients indicated the model was not statistically significant ($p = 0.185$). It correctly classified only 56.3% of cases, showing no improvement over Block 0. Participants with secondary education were 1.349 times more likely to take extra sun precautions (Exp(B) = 1.349, 95% CI [0.951, 1.349]). However, those with post-secondary education had 0.965 times lower odds of taking precautions (Exp(B) = 0.965, 95% CI [0.779, 1.196]). Overall, education level did not significantly contribute to predicting the likelihood of taking extra sun precautions ($p = 0.163$ for completed secondary education, $p = 0.746$ for completed post-secondary education).

Table 2. Binary logistic regression results for participants who answered yes or no to taking extra precautions when there was a high UV rating, categorized by highest level of education (N = 2751). The model predicts how education level affects the likelihood of taking extra sun precautions. The regression coefficient (B) represents the relationship between predictor and outcome variable. The standard error (S.E.) indicates the variability of B from the true value. The Wald test measures the contribution of the predictor to the model. The exponential of B (Exp (B)) reflects the odds ratio for developing the outcome of interest.

	B	S.E.	Wald	Exp (B)	95% CI for Exp (B)		Sig.
					Lower	Upper	
Did not complete secondary	-.308	.070	19.256	.735			<.001
Completed secondary	.124	.089	1.944	1.349	.951	1.349	.163
Completed post-secondary	-.035	.109	.105	.965	.779	1.196	.746

Secondary research question—chi-squared analysis

Table 3 shows the differences in the proportion of individuals taking extra sun precautions (Yes vs. No) between two age groups: younger adults (ages 20–34) and older adults (ages 55 and above). A 2×2 chi-squared analysis showed a 2-sided asymptotic significance of 0.737, with a Yate's continuity correction value of 0.113 at 1 degree of freedom. As it is a value greater than 0.05, there was no significant difference in proportions between the groups.

Table 3. Results of a 2x2 chi-squared analysis based on the proportions of participants who answered yes or no to taking extra precautions when there was a high UV rating and their age group (N = 1534).

	Yes	No	df	Value	Asymptotic Significance (2-sided)
Young adults (20-34) (N = 806)	446	360	1	.113	.737
Older adults (55+) (N = 728)	410	318	1		

Discussion

This study set out to explore whether the highest level of education attained, as a primary analysis, and age, as a secondary analysis, were associated with self-reported sun protection behaviours among Canadians aged 15 and over. Contrary to our expectations and previous research, our regression analysis revealed that education was not a statistically significant predictor of behaviour. Likewise, no significant difference was found between groups of younger adults (20–34 years) and older adults (55+) in their likelihood of taking extra precautions when the UV index is high.

These findings challenge the conclusions drawn by previous studies. For example, Falk and Anderson (2013) found a clear association between higher education levels and greater use of sun protection strategies, attributing this relationship to increased health literacy and access to public health messaging. Similarly, Kalia et al. (2019) suggested that education acts as a key determinant in preventive health behaviour, with those attaining post-secondary education more likely to adopt protective measures. Our results suggest instead that this relationship may not be as straightforward in the Canadian population sampled in 1996, or that education may not translate into behavioural change as strongly as previously assumed. For instance, while individuals with post-secondary education may possess greater awareness of sun safety guidelines, this knowledge does not always lead to consistent action, a conclusion similarly drawn by Almuqati et al. (2019). Behavioural change is complex and can be influenced by a variety of factors beyond knowledge, including personal attitudes, perceived susceptibility to harm, convenience, cultural norms, and trust in public health messaging.

The lack of statistically significant differences in sun protection behaviours across education levels may also be influenced by the nature of the dependent variable in this study. The binary response to whether participants took extra precautions ("yes" or "no") provides only a limited view of protective behaviour, potentially obscuring nuanced behavioural differences across education levels. Prior research often examined specific behaviours such as sunscreen use, wearing protective clothing, or seeking shade, which may capture more subtle differences that a broad yes/no question overlooks (Pinault & Fioletov, 2017). Furthermore, the lack of significance could also be due to education level being a broad variable that may serve as a proxy for more specific constructs like health literacy or access to health information.

The absence of age-related differences in sun protection behaviour is similarly unexpected. Previous literature suggested that older adults are more likely to engage in sun-safe behaviours, possibly due to greater awareness of skin cancer risks or more routine outdoor habits that facilitate protective planning (CDC, 2023; Age and Risk, 2024). However, our results indicated no statistically significant differences in precaution-taking between older and younger adults. One possible explanation is that the impact of public health campaigns and general sun safety awareness was fairly uniform across age groups at the time of data collection, limiting detectable age-related differences. Our findings may suggest that age alone is not a reliable predictor of sun-safe behaviour, at least within the context of this dataset and question format. This challenges assumptions in the existing literature that older adults are consistently more cautious in their sun exposure habits. It also highlights the possibility that other factors such as perceived vulnerability, lifestyle, or occupation may play a more influential role than chronological age in shaping protective behaviours.

This study had several limitations that should be acknowledged. First, the age categories in our dataset were limited, restricting the scope of our secondary question. Initially, we aimed to investigate sun protection behaviours in individuals aged 50 and older, given previous research indicating an increased prevalence of melanoma in this group (“Age and Risk”, 2024). However, the dataset segmented ages into narrower categories (45-54 and 55-64) which forced us to adjust our age criterion to include only individuals 55 and up, which may have excluded relevant participants and limited generalizability. Second, the reliance on self-reported data introduces the potential for recall and social desirability bias. Since the Canadian Sun Exposure Survey was conducted over the phone, respondents may have provided inaccurate or idealized answers, undermining the validity of our findings. Third, the dependent variable used to assess sun protection behaviour was limited in scope. Responses to the question “Do you take extra precautions when UV is high?” were restricted to “yes,” “no,” “do not know,” “refused,” or “not stated.” The absence of more nuanced response options such as “sometimes,” “rarely,” or “usually”, constrained our ability to detect variation in behaviours and may have obscured subtle but meaningful patterns. Fourth, our study was restricted to categorical variables, which limited the range of statistical analyses available. Finally, the data we used were collected in 1996. Given the substantial changes in public awareness, use of social media, and climate-related UV exposure since then, our findings may not reflect current sun protection behaviours among Canadians nor the impact of recent public health campaigns. This gap limits the applicability of our conclusions to present-day contexts.

To build on our findings and address these limitations, future studies should utilize updated, nationally representative datasets that reflect current public health messaging, social media influence, and environmental conditions. Surveys should include more precise age categories and refined response options to better capture behavioural diversity. Conducting web-based surveys with built-in anonymity may also help reduce response bias and improve data accuracy. Additionally, to improve research in this field, surveys should be concise, use neutral language in framing questions, and incorporate techniques such as misleading or redundant items to help identify inconsistent or false responses. Subsequent research could explore a broader array of sun protection behaviours, such as sunscreen use, clothing choices, and time spent outdoors, to better understand specific protective patterns. Mixed methods designs, combining quantitative surveys with qualitative interviews, would provide deeper insight into the psychological and contextual factors that influence sun safety behaviours, such as perceived susceptibility or trust in public health guidance. Future intervention studies should evaluate the effectiveness of targeted campaigns for specific demographic groups, such as younger adults who may underestimate their risk. Longitudinal studies would also be valuable for examining how behaviours evolve overtime in response to education, environmental, or policy changes.

These findings suggest that neither education level nor age alone are reliable predictors of sun protection behaviours, highlighting the need for more nuanced public health strategies. Future skin cancer prevention efforts should go beyond demographic targeting and instead focus on addressing behavioural drivers such as perceived risk, accessibility of protective resources, and trust in health messaging. Tailored interventions that consider these factors, especially in the context of evolving environmental conditions and digital health communication, may be more effective in encouraging consistent sun-safe behaviours across the population.

References

Age and risk. (2024, October 15). AIM at Melanoma Foundation. https://www.aimatmelanoma.org/melanoma-101/understanding-melanoma/melanoma-risk-factors/age-and-risk/

Almuqati, R. R., Alamri, A. S., & Almuqati, N. R. (2019). Knowledge, attitude, and practices toward sun exposure and use of sun protection among non-medical, female, university students in Saudi Arabia: A cross-sectional study. *International Journal of Women's Dermatology*, *5*(2), 105–109. https://doi.org/10.1016/j.ijwd.2018.11.005

Amaro-Ortiz, A., Yan, B., & D'Orazio, J. (2014). Ultraviolet radiation, aging and the skin: Prevention of damage by topical camp manipulation. *Molecules*, *19*(5), 6202–6219. https://doi.org/10.3390/molecules19056202

Barnes, P. W., Williamson, C. E., Lucas, R. M., Robinson, S. A., Madronich, S., Paul, N. D., Bornman, J. F., Bais, A. F., Sulzberger, B., Wilson, S. R., Andrady, A. L., McKenzie, R. L., Neale, P. J., Austin, A. T., Bernhard, G. H., Solomon, K. R., Neale, R. E., Young, P. J., Norval, M., … Zepp, R. G. (2019). Ozone depletion, ultraviolet radiation, climate change and prospects for a sustainable future. *Nature Sustainability*, *2*(7), 569–579. https://doi.org/10.1038/s41893-019-0314-2

Brenner, D. R., Gillis, J., Demers, A. A., Ellison, L. F., Billette, J.-M., Zhang, S. X., Liu, J. L., Woods, R. R., Finley, C., Fitzgerald, N., Saint-Jacques, N., Shack, L., & Turner, D. (2024). Projected estimates of cancer in Canada in 2024. *Canadian Medical Association Journal*, *196*(18). https://doi.org/10.1503/cmaj.240095

Campbell, H. S., & Birdsell, J. M. (1994). Knowledge, beliefs, and Sun Protection Behaviors of alberta adults. *Preventive Medicine*, *23*(2), 160–166. https://doi.org/10.1006/pmed.1994.1022

Canadian Skin Cancer Foundation. (2018). *Skin cancer*. CSCF. https://www.canadianskincancerfoundation.com/skin-cancer/

Centers for Disease Control and Prevention. (2012, May 11). *Sunburn and sun protective behaviors among adults aged 18–29 years - United States, 2000–2010*. https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6118a1.htm

Centers for Disease Control and Prevention. (2023, October 18). *Many older adults don't protect their skin from the sun*. https://www.cdc.gov/cancer/research/many-older-adults-dont-protect-their-skin-from-the-sun.html

Facts & Statistics. (2025, January 21). AIM at Melanoma Foundation. https://www.aimatmelanoma.org/facts-statistics/

Falk, M., & Anderson, C. D. (2013). Influence of age, gender, educational level and self-estimation of skin type on sun exposure habits and readiness to increase Sun Protection. *Cancer Epidemiology*, *37*(2), 127–132. https://doi.org/10.1016/j.canep.2012.12.006

Kalia, S., & Kwong, Y. K. (2019). Relationship between sun safety behaviours and modifiable lifestyle cancer risk factors and Vitamin D Levels. *Photodermatology, Photoimmunology & Photomedicine*, *35*(6), 429–435. https://doi.org/10.1111/phpp.12494

Lee, A., Garbutcheon-Singh, K. B., Dixit, S., Brown, P., & Smith, S. D. (2014). The influence of age and gender in knowledge, behaviors and attitudes towards Sun Protection: A cross-sectional survey of Australian outpatient clinic attendees. *Photodermatology, Photoimmunology & Photomedicine*, *35*(6), 429–435. https://doi.org/10.1007/s40257-014-0106-4

Melanoma stats and facts. (2024, July 19). Melanoma Canada. https://melanomacanada.ca/understanding-skin-cancer/melanoma/stats-and-facts/.

Narayanan, D. L., Saladi, R. N., & Fox, J. L. (2010). Review: Ultraviolet radiation and skin cancer. *International Journal of Dermatology*, *49*(9), 978–986. https://doi.org/10.1111/j.1365-4632.2010.04474.x

Pinault, L., & Fioletov, V. (2017). Sun exposure, sun protection and sunburn among Canadian adults. *Health Reports*, *28*(5), 12–19. https://www.researchgate.net/profile/Vitali-Fioletov2/publication/316994625_Sun_exposure_sun_protection_and_sunburn_among_Canadian_adults/links/591ca4faaca272d31bca988a/Sun-exposure-sun-protection-and-sunburn-among-Canadian-adults.pdf

Purdue, M.P. (2002). Predictors of Sun Protection in Canadian Adults. *Canadian Journal of Public Health*, *93*, 470–474. https://doi.org/10.1007/BF03405042

Raymond-Lezman, J. R., & Riskin, S. (2023). Attitudes, Behaviors, and Risks of Sun Protection to Prevent Skin Cancer Amongst Children, Adolescents, and Adults. *Cureus*, *15*(2), e34934. https://doi.org/10.7759/cureus.34934

Statistics Canada. (1998, May 12). *1996 CENSUS: SOURCES OF INCOME, EARNINGS AND TOTAL INCOME, AND FAMILY INCOME.* Government of Canada. https://publications.gc.ca/collections/collection_2016/statcan/11-001/CS11-001-1998-5-12-eng.pdf

Statistics Canada. (2023). Sun Exposure Survey, 1996 [Canada] [Data file]. Odesi. https://doi.org/10.5683/SP3/JCX7LH

Shoveller, J. A., Lovato, C. Y., Peters, L., & Rivers, J. K. (2000). Canadian National Survey on Sun Exposure & Protective Behaviours: outdoor workers. *Canadian journal of public health, 91*(1), 34–35. https://doi.org/10.1007/BF03404250

Watson, T. P. G., Tong, M., Bailie, J., Ekanayake, K., & Bailie, R. S. (2024a). Relationship between climate change and skin cancer and implications for prevention and management: A scoping review. *Public Health*, *227*, 243–249. https://doi.org/10.1016/j.puhe.2023.12.003

Investigating the Health Outcomes of Adhering to the Fruit and Vegetable Consumption Recommended by Canada's Food Guide Before and After the 2019 Revision

Elizabeth Noh, Saikeerthana Sundar, and Tetyana Revchuk

School of Interdisciplinary Sciences, McMaster University

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8th, 2025

Introduction

Canada's Food Guide (CFG) is a resource for healthy eating (Lavigne & Lengyel, 2019). Health Canada periodically revises CFG to address dietary challenges and ensure it aligns with the latest evidence on sodium, saturated fat, and sugar consumption (Health Canada, 2019).

After 2007, the latest version of CFG was released in 2019, a model that promises a holistic nutritional approach, focusing on eating habit consciousness (Lavigne & Lengyel, 2019). Unlike the previous version, which provided serving numbers and portion size recommendations, it uses visual depictions for the proportions of the food categories (Webster, 2019). The significant changes to the CFG include emphasizing plant-based consumption, decreasing the intake of processed meats and foods containing saturated fats, and making water the primary drink of choice. While many public health and nutrition experts favour the new model, others are skeptical, believing the new representation may be too vague.

The 2019 CFG recommends that fruits and vegetables make up half of your plate per meal (Lavigne & Lengyel, 2019). Fruits and vegetables have long been considered healthy and are rich sources of vitamins, minerals, phytochemicals and dietary fiber (Slavin & Lloyd, 2012). These compounds, as well as the energy density and low glycemic load found in these foods, are thought to be involved in protective mechanisms (Park, 2021; Aune et al., 2017). Studies have shown that the consumption of such foods is associated with decreased risks of type 2 diabetes, cancer, cardiovascular disease, stroke and all-cause mortality (Aune et al., 2017; Madsen et al., 2023; Wang et al., 2015). Therefore, a change in Canadians' consumption of fruits and vegetables could also be accompanied by a change in health-related outcomes such as blood pressure (BP), blood cholesterol/lipids, and Body Mass Indexes (BMIs).

High BP (hypertension) is a condition in which the pressure exerted by blood on the walls of the arteries is above normative levels, possibly leading to adverse effects such as increased risk of cardiovascular disease (Oparil et al., 2018). In Canada, it is considered high risk if blood pressure is 135+/85+ mmHG (Heart&Stroke, n.d.). The correlation between an elevated consumption of fruits and vegetables and reduced BP has been inconsistent across research. However, a meta-analysis revealed that this correlation is robust across several studies (Madsen et al., 2023).

As per the National Heart, Lung, and Blood Institute (NHLBI, 2024), the buildup of cholesterol-transporting molecules in the arteries can lead to heart attacks or strokes. According to Statistics Canada (2013), a total cholesterol level over 6.2 mmol/L is considered dangerous. The relationship between blood cholesterol and fruit and vegetable consumption is conflicting across studies. However, a study by Schoeneck & Iggman (2021) shows that the intake of fruits and vegetables is associated with decreased blood cholesterol, with the size of reduction varying across different foods.

Body Mass Index (BMI) is an estimate of body fat based on height and weight, with a higher BMI associated with an increased body fat percentage (Cleveland Clinic, 2024). Elevated body fat can have adverse effects such as increased risk of cardiovascular disease, increased BP, diabetes, and stroke (Harvard Health, 2017). Studies have shown that an increased consumption of fruits and vegetables can be associated with normative BMIs as the high fiber and low-calorie content of fruits and vegetables increases satiation (Nour et al., 2018).

Several articles have examined the awareness, knowledge, and use of CFG and found it to be generally higher for the 2019 version compared to the one released in 2007 (Barco Leme et al., 2021; Ramuscak et al., 2022). However, a knowledge gap remains on whether Canadians are eating healthier and if any significant health effects have resulted from following the 2019 CFG.

The Canadian Community Health Survey (CCHS) is an annual questionnaire which collects health data, including information regarding participants' adherence to the CFG and dietary habits (Statistics Canada, 2024). To address the knowledge gap mentioned, CCHS data can be accessed to determine if the revision of CFG in 2019 could have contributed to a change in the number of fruits and vegetables consumed by Canadians who follow CFG. Fruit and vegetable consumption could be followed to track adherence to the proposed focus on plant-based consumption. CCHS data from 2017-2018 and 2019-2020 are available, allowing the examination of the above-mentioned relationship before and after the 2019 revision.

Our study aims to bridge gaps in the current literature by understanding eating habits and health outcomes after the 2019 CFG revision. It raises the question: Is there a difference in the self-reported average daily consumption of fruit and vegetables (recorded as number of times per day) in Canadians aged 18-49 who follow Canada's Food Guide (self-reported) in 2019-2020 compared to Canadians of the same age in 2017-2018?

To further investigate if the CFG revisions, particularly in fruits and vegetables consumption, could have contributed to a change in BP, blood cholesterol/lipids and BMI, the consequent follow-up research questions could be asked: Is there a difference in the prevalence of self-reported high blood pressure (primary outcome), high blood cholesterol/lipids (primary outcome), and an overweight/obese BMI (secondary outcome) in Canadians, aged 18 to 49, between 2017-2018 and 2019-2020? Does the self-reported average daily consumption of fruits and vegetables correlate with the blood pressure, blood cholesterol/lipids, and BMI in 2019-2020 in Canadians aged 18-49?

Methods

Study Design and Database

The CCHS is a cross-sectional, observational study collecting health data on Canadians since 2001. The questionnaire surveys Canadians from January of one year till December of the next, and databases released are consistent in their "design and population representation" so that they can be compared to one another (Government of Canada, 2023). The CCHS databases from 2017 to 2018 and 2019 to 2020 were accessed from the Canadian research data repository Borealis (https://borealisdata.ca). Data from both databases was used to answer the research questions. The main objective of the CCHS is to gather health data and surveillance for the Canadian population, informing plans and policies to improve national health. Researchers collected self-reported data from participants through telephone and personal interviews (Government of Canada, 2023).

Study Participants

A total of 255 participants between the ages of 18 and 49 who indicated that they used the CFG "to plan meals/grocery shopping" were included in this study. All those who did not indicate this were considered not following the CFG. This age group was chosen to close a research gap as previous studies about the CFG have focused on young children and their parents (Barco Leme et al., 2021; Ramuscak et al., 2022). Participants over the age of 49 were not involved as risk factors like high BP and cholesterol tend to advance with older age (Singh et al., 2012), which may skew results. The use of 2017-2018 and 2019-2020 CCHS datasets allowed for a comparison of output data from participants actively using CFG before and after the 2019 revision. Participants with the following answers recorded for the categories regarding self-reported BMI, high BP, high blood cholesterol/lipids and the average daily consumption of fruits and vegetables were excluded from the study: valid skip, refusal, don't know, not stated.

Study Outcomes

The primary outcome of this study (question 1) is the average daily consumption of fruits and vegetables. Participants reported the number of fruits and vegetables they consumed as the number of times per day, week or month. Fruits and vegetables were defined as belonging to various categories, such as fruit juices, solid fruits, green vegetables, orange vegetables, etc. (Statistics Canada, 2020b). Researchers used these results to calculate the average daily consumption of fruits and vegetables, which was recorded as continuous data.

The presence of high BP (yes or no) and high cholesterol/lipids (yes or no) was assessed with the questions "do you have high blood pressure?" and "do you have high blood cholesterol or lipids?", respectively (Statistics Canada, 2020a).

Researchers calculated BMI based on self-reported height and weight. Participants were then categorized as underweight/normal weight or overweight/obese class I, II, or III according to international standards set by the World Health Organization and Health Canada (Statistics Canada, 2020b).

It is not relevant to examine both the 2017–2018 and 2019-2020 fruits and vegetable consumption correlations with selected health outcomes to determine the effects of the 2019

CFG revisions. This would be redundant as our first and second research questions compared the differences in fruit and vegetable consumption and health outcomes between the selected years.

Additional Data Collection

Demographic data, including the participants' sex and age, was gathered to describe the sample population. Participants selected their sex (male or female) and their age value.

Statistical Analysis

The IBM Statistical Package for Social Sciences (Version 27) was used to conduct statistical tests. Demographic data and categorical variables (BP, blood cholesterol/lipids, and BMI) were represented as frequency (n) and percentages (%). Continuous data (average daily consumption of fruits and vegetables) was reported as median ± inter quartile range (IQR). A preliminary comparison between the proportion of people who followed the CFG before and after the revision was examined using a Chi-squared test. To assess the normality of average daily consumption, the Kolmogorov-Smirnov test was performed. For question 1, the Mann-Whitney U test was used to compare the average daily consumption of fruits and vegetables, as the data was not normally distributed. To compare the prevalence of BP and blood cholesterol/lipids before and after the revision, a Chi-squared test was used to answer question 2. A Fisher's exact test was used for BMI, as expected and observed values were <5. To study the relationship between the average daily consumption of fruits and vegetables, BP, blood cholesterol/lipids, and BMI post revision, Spearman's rho was used for question 3. A p-value (P) < 0.05 was considered statistically significant for all tests.

Results

The CCHS databases contained 113,290 participants in 2017-2018 and 108,252 participants in 2019-2020, of whom 178 and 4,667 participants used CFG to plan meals/grocery shopping, respectively**.** A Chi-squared test showed a statistically significant difference in the number of participants who used CFG to plan meals/grocery shopping between the 2017-2018 and 2019-2020 groups at a $P<0.00001$.

Participants with missing data for one or more dependent variables were excluded, leading to 73 participants from 2017-2018 and 182 participants from 2019-2020 being included in this study. In the 2017-2018 and 2019-2020 databases, 462 and 380 participants were missing data reporting if they had high blood pressure, 10068 and 7753 were missing data reporting if they had high blood cholesterol/lipids, and 15297 and 6499 participants were missing data regarding their BMI classification, respectively.

In the 2017-2018 data, 48 (65.7 %) participants were 18-34 years old, and 25 (34.3 %) were 35-49. In the 2019-2020 data, 75 (41.2 %) were 18-34 years old, and 25 (58.8 %) were 35-49 years old. The median consumption of fruits and vegetables was 3.9 (IQR=2.3) and 3.7 (IQR=3.0) for 2017-2018 and 2019-2020, respectively (**Table 1**)**.**

Table 1: Health and descriptive characteristics of Canadians part of the CCHS database who followed Canada's Food Guide to Make Dietary Decisions in 2017-2018 and 2019-2020 [N=255].

		Canadian Community Health Survey	
		2017-2018 (n = 73)	2019-2020 (n = 182)
Sex	Male	26 (35.6)	52 (28.6)
	Female	47 (64.4)	130 (71.4)
Age	18-34	48 (65.7)	75 (41.2)
	35-49	25 (34.3)	107 (58.8)
Average daily consumption of fruits & vegetables		3.9 ± 2.3	3.7 ± 3.0
High blood pressure	Yes	8 (11)	17(9.3)
	No	65 (89)	165(90.7)
High cholesterol	Yes	1 (1.4)	9(4.9)
	No	72 (98.6)	173(95.1)
Body mass index (BMI)	Underweight/normal	34 (46.6)	69 (37.9)
	Overweight/obese	39 (53.4)	113(62.1)

Table 1 is based on data from the Canadian Community Health Survey (CCHS) for the years 2017-2018 and 2019-2020 and only includes Canadians who followed Canada's Food Guide (CFG) to plan meals/grocery shopping. Participants with missing data for the continuous or categorical variables included in Table 1 were excluded. Categorical data is expressed as a frequency (n) followed by a percentage (%). The average daily consumption of fruits and vegetables (continuous) is reported as median ± interquartile range (IQR).

A Kolmogorov-Smirnov test revealed that neither of the annual datasets was normally distributed for the average daily consumption of fruits and vegetables (**Figure 1**)**.** A Mann-Whitney U test revealed no statistically significant difference in the median self-reported average daily consumption of fruits and vegetables among Canadians aged 18-49 who followed Canada's Food Guide in 2017-2018 compared to 2019-2020 (**Figure 1**).

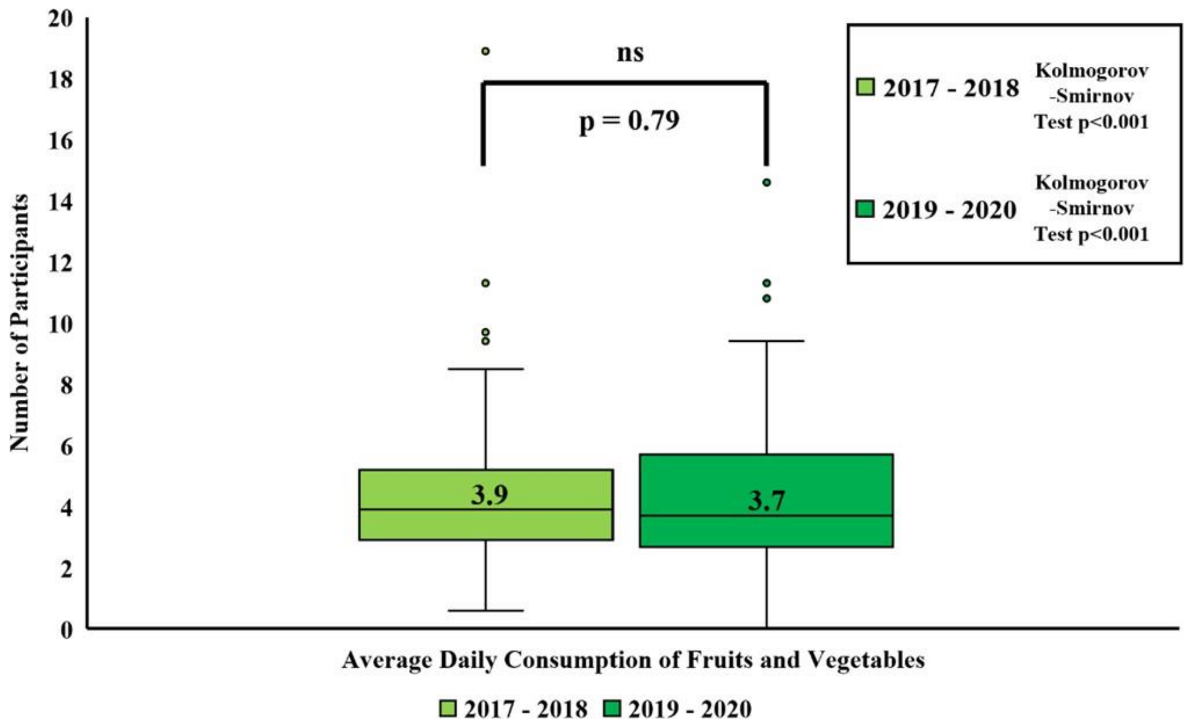


Figure 1. The Distribution of Participants' Average Daily Consumption of Fruits and Vegetables in Canadians Aged 18-49 Who Followed Canada's Food Guide in 2017-2018 Compared to Canadians of the Same Age in 2019-2020 (n= 73 for 2017-2018 and n=182 for 2019-2020). Participants self–reported whether they used CFG to plan their meals/grocery shopping and their average daily consumption of fruits and vegetables.

Investigation into this study's second research question revealed no statistically significant differences in outcomes related to high BP, high blood cholesterol/lipids, and an overweight/obese BMI between the 2017-2018 sample and the 2019-2020 sample. In the 2017-2018 and 2019-2020 samples, 8 (11 %) and 17(9.3%) participants had high BP, respectively. A Chi-squared test revealed no statistically significant difference in the number of participants with and without high BP between the 2017-2018 and 2019-2020 groups (**Figure 2A**)**.** No statistical difference was found in the prevalence of high blood cholesterol/lipids between 2017-2018 and 2019-2020 using a Fisher's Exact Test, and 1 (1.4 %) and 9 (4.9 %) of participants had high blood cholesterol/lipids, respectively (**Figure 2B**). In the 2017-2018 and 2019-2020 samples, 39 (53.4 %) and 113(62.1%) were overweight/obese, respectively. Using a Chi-squared test, no statistical difference was found in the prevalence of an overweight/obese BMI between the two groups **(Figure 2C**)**.**

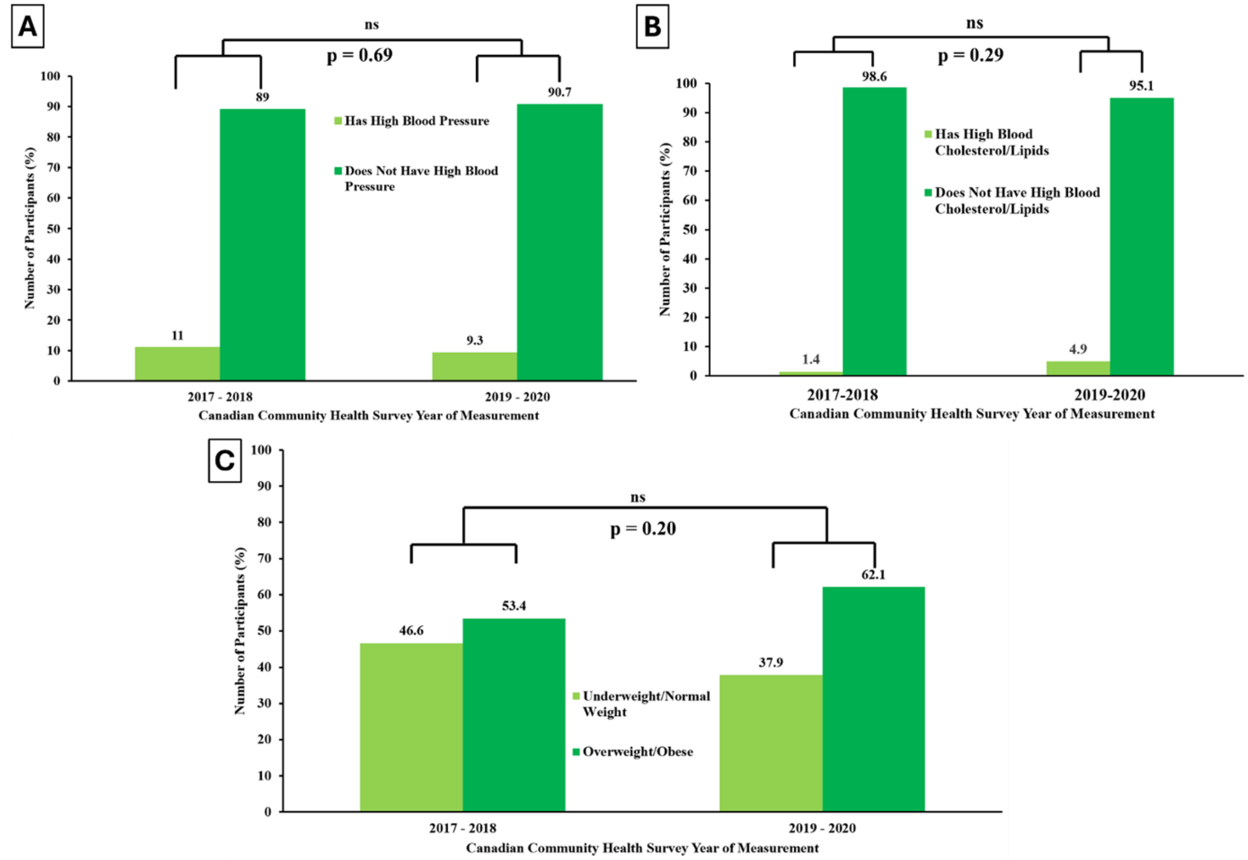


Figure 2. The Prevalence of High Blood Pressure, High Blood Cholesterol/Lipids, and Overweight/Obese BMI Classification Amongst Canadians Aged 18-49 Who Followed Canada's Food Guide in the Years 2017-2018 and 2019-2020 (n= 73 for 2017-2018 and n=182 for 2019-2020). The number of participants is a % respective to the annual dataset to which they belong. **A** - Participants self-reported whether they had high blood pressure or not. **B** - Participants self-reported if they had high blood cholesterol/lipids or not. **C** - Participants' BMI was calculated based on their height and weight and grouped as underweight/normal or overweight/obese.

A Spearman-Rho correlation analysis was conducted to investigate the third research question and determine whether the average daily consumption of fruits and vegetables in 2019-2020 correlates with the prevalence of high BP, high blood cholesterol/lipids, and an overweight/obese BMI. A Spearman-Rho correlation test revealed that none of the three dependent variables mentioned correlate with self-reported average daily consumption of fruits and vegetables (**Table 2**).

Table 4: Spearman-Rho Correlation Analysis Between the Average Daily Consumption of Fruits and Vegetables and Blood Pressure, Blood Cholesterol/Lipids, and BMI of Canadians Aged 18-49 Who Followed CFG to plan meals/grocery shopping in 2019-2020 [N=182].

		Average Daily Consumption of Fruits and Vegetables	BMI	Blood Pressure	Blood Cholesterol /Lipids
Average Daily Consumption of Fruits and Vegetables	Correlation Coefficient	1.000	-0.059	0.032	-0.096
	Sig. (2-tailed)	-	0.427	0.670	0.196
	N	182	182	182	182

Discussion

This study investigated whether the consumption of fruits and vegetables changed in Canadians aged 18-49 following CFG before and after the 2019 revision and if this influenced key health outcomes. When investigating the first research question, we found no significant differences in fruit and vegetable consumption between 2017-2018 and 2019-2020. This suggests that Canadians of this age group did not change their fruit and vegetable eating habits after the revision. These results could be attributed to financial stress, the rise in food and ingredient prices over the past few years, and the ambiguity of the new food guide. Recent studies have also shown that compared to the 2007 version of the CFG, the latest version is more expensive for adults to adhere to (Taylor et al., 2023). Furthermore, studies state that the pictorial depictions of the food portions in the revised CFG could offer vague guidance to the readers, resulting in difficulty following it (Webster, 2019).

Moreover, we found no statistically significant difference in self-reported high BP, high blood cholesterol/lipids, and an overweight/obese BMI in Canadians aged 18-49 between 2017-2018 and 2019-2022 when investigating the second research question. This suggests no short-term changes in these health outcomes due to following the new food guide. Previous studies have shown that dietary patterns that include higher proportions of plant-based foods positively affect BP, cholesterol, cardiovascular disease, cancer, and more (Cena & Calder, 2020). As we did not see a change in fruit and vegetable consumption after the revision, the lack of change in health outcomes could be due to the Canadian population not increasing their fruit and vegetable consumption. Alternatively, there could be other variables that confound our results. To gain more insight into this, these factors should be further explored in future studies.

Furthermore, we have found no correlation between the self-reported average daily consumption of fruits and vegetables with BP, blood cholesterol/lipids, and BMI in 2019-2020 in Canadians aged 18-49 when investigating the third research question. This finding suggests that the fruit and vegetable consumption in this group of Canadians alone may not be indicative of their health outcomes and that there could be other factors at play. This contradicts a meta-analysis of 95 studies that showed a reduced risk of cardiovascular disease and all-cause mortality due to increased fruit and vegetable consumption (Aune et al., 2017). The results from Aune et al. (2017), suggest that dietary guideline revisions alone may not be sufficient to drive measurable health changes in the short term, and more research should be done on the consumption of other food groups and their contribution to overall health.

A study that conducted a cross-sectional analysis to determine the association between dietary quality and dietary guideline adherence with mental health outcomes has also found that there is a significant association between dietary quality and well-being but no association between guideline adherence and well-being (Meegan et al., 2017). This is consistent with our findings. However, the study by Megan et al. (2017) primarily looked at mental health outcomes, which were not evaluated in our study. Moreover, our results are consistent with other studies done in America. For example, one study has stated that there is no comparable mandate to evaluate the effectiveness of the dietary guidelines for Americans at periodic intervals (Guthrie & Smallwood, 2003), which aligns with our findings.

On the contrary, a recent study examined the influence of adhering to the 2019 CFG's recommendations on cardiovascular disease. The results suggested that greater adherence to the healthy choices recommended by the new CFG reduced cardiovascular disease by 24 % in middle-aged and older adults (Brassard et al., 2022). The results of our study could differ from those of Brassard et al. (2022) due to some of the weaknesses found in our study.

Although there was a large sample size, all the data used was retrospective and came from a self-report questionnaire, which is liable to response biases (Rosenman et al., 2011). To calculate food consumption, the participants were asked about their eating habits over the past month, and then these values were generalized to produce the average annual consumption. While this is a viable method to report on previous eating patterns, it is not the most robust. First, there is a potential for memory inaccuracies as past events are reported. Second, we cannot be sure that the nutritional intake during the questionnaire is representative of the entire year. A future longitudinal study could follow a group of people throughout the year and have them report their food intake after each meal. BP and blood cholesterol/lipids could be measured at a clinic at different times during the year to provide numerical data instead of a "yes" or "no" answer. The BMI in the survey was also clustered into groups, so in the following study, raw values could be obtained to study how they correlate with the consumption of fruits and vegetables. These changes will limit the use of self-reported data, where possible, while setting other parameters to limit potential biases.

Our study suggests no difference in dietary and health changes in Canadians aged 18-49 who follow the CFG before and after the 2019 revision. As a next step, it would be interesting to investigate whether there is a difference in these factors between those who follow the CFG and those who do not, as well as within different age groups. Our study only looked at the short-term dietary and health outcomes directly following the recent CFG revision. Future studies should examine if there are any long-term alterations associated with the new CFG. Potential confounding factors and covariates, such as socioeconomic status, which has been implicated in food consumption and health outcomes (Pechey & Monsivais et al., 2016; St John et al., 2008), and the consumption of other food groups should also be explored.

While the 2019 revision of Canada's Food Guide aimed to encourage healthier eating habits, our findings suggest that guideline updates alone may not lead to meaningful dietary changes without complementary public health initiatives. These results can be considered when developing new strategies to support long-term dietary adherence. For example, studies have shown that lower income groups adhere less to the guidelines concerning the consumption of fruits compared to those belonging to higher income groups (Dijkstra et al., 2014). Additionally, studies indicate that current diets are largely environmentally unsustainable, with meat and dairy products being the primary contributors to dietary greenhouse gas emissions. Therefore, shifting toward a plant-based diet and limited animal food improves overall diet quality and environmental sustainability (Trolle et al., 2024). The Canadian government can take these findings and other studies into consideration to promote individuals to follow the CFG, which will lead to meaningful dietary changes for health initiatives as well as a sustainable environment. An example of this notion could be guideline changes, bringing back more specific recommendations regarding serving sizes. Revising dietary guidelines is important in promoting public health, but additional strategies may be required to ensure meaningful behavioural change and measurable health benefits.

References

Aune, D., Giovannucci, E., Boffetta, P., Fadnes, L. T., Keum, N., Norat, T., Greenwood, D. C., Riboli, E., Vatten, L. J., & Tonstad, S. (2017). Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality—A systematic review and dose-response meta-analysis of prospective studies. International Journal of Epidemiology, 46(3), 1029–1056. https://doi.org/10.1093/ije/dyw319

Barco Leme, A. C., Laila, A., Hou, S., Fisberg, R. M., Ma, D. W. L., Fisberg, M., & Haines, J. (2021). Perceptions of the 2019 Canada’s food guide: A qualitative study with parents from Southwestern Ontario. Applied Physiology, Nutrition, and Metabolism, 47(1), 34–40. https://doi.org/10.1139/apnm-2021-0414

Brassard, D., Manikpurage, H. D., Thériault, S., Arsenault, B. J., & Lamarche, B. (2022). Greater adherence to the 2019 Canada’s Food Guide recommendations on Healthy Food Choices reduces the risk of cardiovascular disease in adults: A prospective analysis of UK Biobank Data. The American Journal of Clinical Nutrition, 116(6), 1748–1758. https://doi.org/10.1093/ajcn/nqac256

Cena, H., & Calder, P. C. (2020). Defining a healthy diet: Evidence for the role of contemporary dietary patterns in health and disease. Nutrients, 12(2), 334. https://doi.org/10.3390/nu12020334

Cleveland Clinic. (2024, October 28). Body mass Index (BMI). https://my.clevelandclinic.org/health/articles/9464-body-mass-index-bmi

Dijkstra, S. C., Neter, J. E., Brouwer, I. A., Huisman, M., & Visser, M. (2014). Adherence to dietary guidelines for fruit, vegetables and fish among older Dutch adults; the role of education, income and job prestige. The Journal of Nutrition, Health and Aging, 18(2), 115–121. https://doi.org/10.1007/s12603-013-0402-3

Government of Canada, S. C. (2023, December 29). Canadian Community Health Survey—Annual Component (CCHS). https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3226

Guthrie, J. F., & Smallwood, D. M. (2003). Evaluating the effects of the dietary guidelines for Americans on consumer behavior and Health: Methodological Challenges. Journal of the American Dietetic Association, 103(12), 42–49. https://doi.org/10.1016/j.jada.2003.09.036

Harvard Health. (2017, January 20). Abdominal obesity and your health. https://www.health.harvard.edu/staying-healthy/abdominal-obesity-and-your-health#:~:text=Excess%20body%20fat%20has%20serious,information%20from%20Harvard%20Medical%20School.&text=I'd%20like%20to%20receive,for%20only%20$4.99%20a%20month.&text=Already%20a%20member?,Login%20%C2%BB.

Health Canada. (2019). Canada-Food-Guide-Presentation-Eng. Canada. https://www.canada.ca/content/dam/hc-sc/documents/services/canada-food-guide/resources/stakeholder-toolkit/canada-food-guide-presentation-eng.pdf

Heart&Stroke. (n.d.). High blood pressure. Heart and Stroke Foundation of Canada. https://www.heartandstroke.ca/heart-disease/risk-and-prevention/condition-risk-factors/high-blood-pressure?gad_source=1&gclid=CjwKCAiAiaC-BhBEEiwAjY99qCBZ4FyJcLQpqGtPSE43W8FbgqVMNFe8Fn2svAKMINZEUtzzw784rhoCqGoQAvD_BwE&gclsrc=aw.ds

Lavigne, S. E., & Lengyel, C. (2019). A new evolution of Canada's Food Guide. Canadian Journal of Dental Hygiene, 53(3), 143–145.

Madsen, H., Sen, A., & Aune, D. (2023). Fruit and vegetable consumption and the risk of hypertension: A systematic review and meta-analysis of prospective studies. European Journal of Nutrition, 62(5), 1941–1955. https://doi.org/10.1007/s00394-023-03145-5

Meegan, A., Perry, I., & Phillips, C. (2017). The association between dietary quality and dietary guideline adherence with mental health outcomes in adults: A cross-sectional analysis. Nutrients, 9(3), 238. https://doi.org/10.3390/nu9030238

National Heart, Lung, and Blood Institute. (2024, April 17). What is Blood Cholesterol? NHLBI. https://www.nhlbi.nih.gov/health/blood-cholesterol
Nour, M., Lutze, S., Grech, A., & Allman-Farinelli, M. (2018). The Relationship between Vegetable Intake and Weight Outcomes: A Systematic Review of Cohort Studies. Nutrients, 10(11), 1626. https://doi.org/10.3390/nu10111626

Oparil, S., Acelajado, M. C., Bakris, G. L., Berlowitz, D. R., Cífková, R., Dominiczak, A. F., Grassi, G., Jordan, J., Poulter, N. R., Rodgers, A., & Whelton, P. K. (2018). Hypertension. Nature reviews. Disease primers, 4, 18014. https://doi.org/10.1038/nrdp.2018.14

Park, H. A. (2021). Fruit Intake to Prevent and Control Hypertension and Diabetes. Korean Journal of Family Medicine, 42(1), 9–16. https://doi.org/10.4082/kjfm.20.0225

Pechey, R., & Monsivais, P. (2016). Socioeconomic inequalities in the healthiness of food choices: Exploring the contributions of food expenditures. Preventive Medicine, 88, 203–209. https://doi.org/10.1016/j.ypmed.2016.04.012

Ramuscak, A. V., Ma, D. W., Forbes, L. E., Duncan, A. M., Sadowski, A., & Haines, J. (2022). Comparing the perceptions and opinions of the 2007 and 2019 Canada's food guides among parents of young children. Frontiers in Public Health, 10. https://doi.org/10.3389/fpubh.2022.944648

Rosenman, R., Tennekoon, V., & Hill, L. G. (2011). Measuring bias in self-reported data. International Journal of Behavioural and Healthcare Research, 2(4), 320. https://doi.org/10.1504/ijbhr.2011.043414

Schoeneck, M., & Iggman, D. (2021). The effects of foods on LDL cholesterol levels: A systematic review of the accumulated evidence from systematic reviews and meta-analyses of randomized controlled trials. Nutrition Metabolism and Cardiovascular Diseases, 31(5), 1325–1338. https://doi.org/10.1016/j.numecd.2020.12.032

Singh, G. M., Danaei, G., Pelizzari, P. M., Lin, J. K., Cowan, M. J., Stevens, G. A., Farzadfar, F., Khang, Y.-H., Lu, Y., Riley, L. M., Lim, S. S., & Ezzati, M. (2012). The age associations of blood pressure, cholesterol, and glucose. Circulation, 125(18), 2204–2211. https://doi.org/10.1161/circulationaha.111.058834

Slavin, J. L., & Lloyd, B. (2012). Health benefits of fruits and vegetables. Advances in Nutrition, 3(4), 506–516. https://doi.org/10.3945/an.112.002154

Statistics Canada. (2020a) Canadian community health survey (CCHS) annual component – 2017-2018 questionnaire. https://hdl.handle.net/11272.1/AB2/SEB16A/LGXDPU

Statistics Canada. (2020b) Canadian community health survey (CCHS) annual component, PUMF, 2017-2018 derived variable (DV) specifications. https://hdl.handle.net/11272.1/AB2/SEB16A/WDZZUQ

Statistics Canada (2023) Canadian community health survey, 2017-2018: Annual component (V2). Borealis. https://doi.org/10.5683/SP3/EYLZ18

Statistics Canada (2024) Canadian community health survey, 2019-2020: Annual component (V3). Borealis. https://doi.org/10.5683/SP3/ZVCGBK

Statistics Canada (2024) Canadian community health survey (CCHS) annual component 2019-2020 - PUMF derived variable (DV) specifications. https://hdl.handle.net/11272.1/AB2/DMJPVY/CW2GZ3

Statistics Canada (2013) Cholesterol levels of Canadians, 2009 to 2011 https://www150.statcan.gc.ca/n1/pub/82-625-x/2012001/article/11732-eng.htm

St John, M., Durant, M., Campagna, P. D., Rehman, L. A., Thompson, A. M., Wadsworth, L. A., & Murphy, R. J. (2008). Overweight Nova Scotia children and youth: The roles of household income and adherence to Canada's Food Guide to Healthy Eating. Canadian Journal of Public Health, 99(4), 301–306. https://doi.org/10.1007/bf03403760

Taylor, S., Charlebois, S., & Music, J. (2023). Affordability of Canada's Food Guide: Current challenges amid covid-19, war in Ukraine, and other world events. Frontiers in Nutrition, 10. https://doi.org/10.3389/fnut.2023.1085855

Trolle, E., Meinilä, J., Eneroth, H., Meltzer, H. M., Þórsdóttir, I., Halldorsson, T., & Erkkola, M. (2024). Integrating environmental sustainability into food-based dietary guidelines in the Nordic countries. Food & Nutrition Research, 68. https://doi.org/10.29219/fnr.v68.10792

Wang, P., Fang, J., Gao, Z., Zhang, C., & Xie, S. (2015). Higher intake of fruits, vegetables or their fiber reduces the risk of type 2 diabetes: A meta‐analysis. Journal of Diabetes Investigation, 7(1), 56–69. https://doi.org/10.1111/jdi.12376

Webster, P. (2019). Canada's updated food guide promotes mindful eating. The Lancet (British Edition), 393(10170), e5–e5. https://doi.org/10.1016/S0140-6736(19)30202-8

National Heart, Lung, and Blood Institute. (2024, April 17). What is Blood Cholesterol? NHLBI. https://www.nhlbi.nih.gov/health/blood-cholesterol

The Impact of Early Tobacco Use on Self-Reported Anxiety and Depression: Examining the Role of Sex

Mohammed Al-Rawi, Laveenah Jeyapalan, Nilakshi Rajmogan

Department of Life Sciences, McMaster University

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8, 2025

Introduction

Mental health is a critical issue that continues to require significant societal attention. The rise of social media and technology has influenced mental health in both positive and negative ways, contributing to its increasing presence in individuals. While many mental health disorders stem from uncontrollable factors, certain lifestyle choices can also play a role in their development. One such factor is tobacco use, which has been shown to negatively impact the development of nervous structures, cognitive functions, and overall increasing the susceptibility of mood disorders (Hajdusianek et. al, 2021).

Although short-term nicotine use can temporarily enhance cognition, learning, and memory, chronic tobacco smoking has the opposite effect, leading to cognitive deficits, impaired verbal memory, and slower processing speed (Valentine & Sofuoglu, 2018). Research indicates that smoking heightens anxiety and tension, despite the common perception that it relieves stress (Mental Health Foundation, 2021). Nicotine affects the brain's neurotransmitter systems by altering dopamine release and other chemicals involved in mood regulation (Lutfy et.al, 2013). In addition to these neurochemical effects, nicotine is a strong activator of the hypothalamus pituitary adrenal (HPA) axis, a major component of the body's stress response system, due to its release of hormones such as cortisol (Lutfy et.al, 2013). This activation can disrupt the body's natural ability to handle stress (Wu, 2023). Hence, while it may provide a temporary sense of relaxation, this effect is misleading, as nicotine dependence leads to withdrawal symptoms such as irritability, anxiety, and heightened stress. This may increase overall stress and anxiety over time, as the body becomes more dependent on nicotine rather than its natural stress-regulation mechanisms (Wu, 2023).

On a neurological level, structural neuroimaging studies have linked prolonged tobacco use to cortical thinning and a reduction in the size of various brain structures (Elbejjani et al., 2019). The age of tobacco initiation further influences these effects (Dai et.al, 2022). Beyond biological vulnerability, early tobacco use may influence long-term psychological well-being. Adolescence is a crucial period for emotional development, and behaviors adopted during this stage can shape mental health outcomes later in life (World Health Organization, 2024). Tobacco impacts the prefrontal cortex, the region responsible for decision-making, planning, and impulse control (Goriounca & Mansvelder, 2012). Since the prefrontal cortex is the last part of the brain to fully develop, adolescents are particularly vulnerable. This heightened sensitivity makes them more prone to risky behaviours and increases their likelihood of developing nicotine addiction if they begin smoking at an early age (Goriounca & Mansvelder, 2012).

Studies have established a connection between tobacco use and the prevalence of mental health disorders; however, there is limited research on investigating how the age of tobacco initiation influences the prevalence of self-reported anxiety and depression (Fluharty, 2017). Some studies have shown that early exposure to nicotine can increase the risk of mental health issues by altering brain development, however; very few studies have directly compared rates of anxiety and depression between individuals who start using tobacco at a young age and those who start later. (Laviolette, 2021).

Additionally, while there are many studies that demonstrate the sex differences in both tobacco use patterns and mental health outcomes, there is a gap in research delving into how sex moderates the relationship between early tobacco use and self-reported mental health disorders (Thompson et.al, 2016). It has generally been found that women initiate smoking at later ages, yet present more negative health outcomes than men (Thomspon et.al, 2016). Furthermore, women are statistically more likely to report symptoms of anxiety and depression, which may interact with tobacco use in unique ways (Shi, 2021). Most studies in current literature examine smoking behavior in general and do not explore how early initiation affects men and women differently in terms of psychological distress. The relationship between sex, age of tobacco initiation, and mental health outcomes remains an under explored but crucial area of inquiry.

This study aims to address this gap by analyzing the difference in the prevalence of self-reported anxiety and depression among individuals who initiate tobacco use early (youth aged 17 and younger) compared to those who start later. To follow-up, this study will also be addressing if sex moderates this relationship, specifically whether early tobacco use affects self-reported anxiety and depression differently in males and females. By addressing this gap, the findings could contribute to a deeper understanding of how the timing of tobacco use onset and sex influence mental health outcomes, potentially informing targeted public health interventions and prevention strategies aimed at adolescents and young adults.

Methods

This study used a cross-sectional design to examine the relationship between early tobacco initiation and self-reported symptoms of anxiety and depression, with a focus on potential sex differences. This data analysis used the 2022 Mental Health and Access to Care Survey (MHACS), accessed through Odesi. The MHACS was designed by Statistics Canada as a cross-sectional, nationally representative survey to assess mental health outcomes and access to care across Canada (Stephenson, 2023).

Participants included 1,132 individuals between the ages of 15 and 74 who reported their history of tobacco use, anxiety, and depression. Inclusion criteria required participants to have complete data on tobacco use initiation, anxiety and depression symptoms, and sex. Participants were excluded if they had missing data on any of these variables or reported other substance use at the time of tobacco initiation to limit potential confounding effects (**Figure 1**).

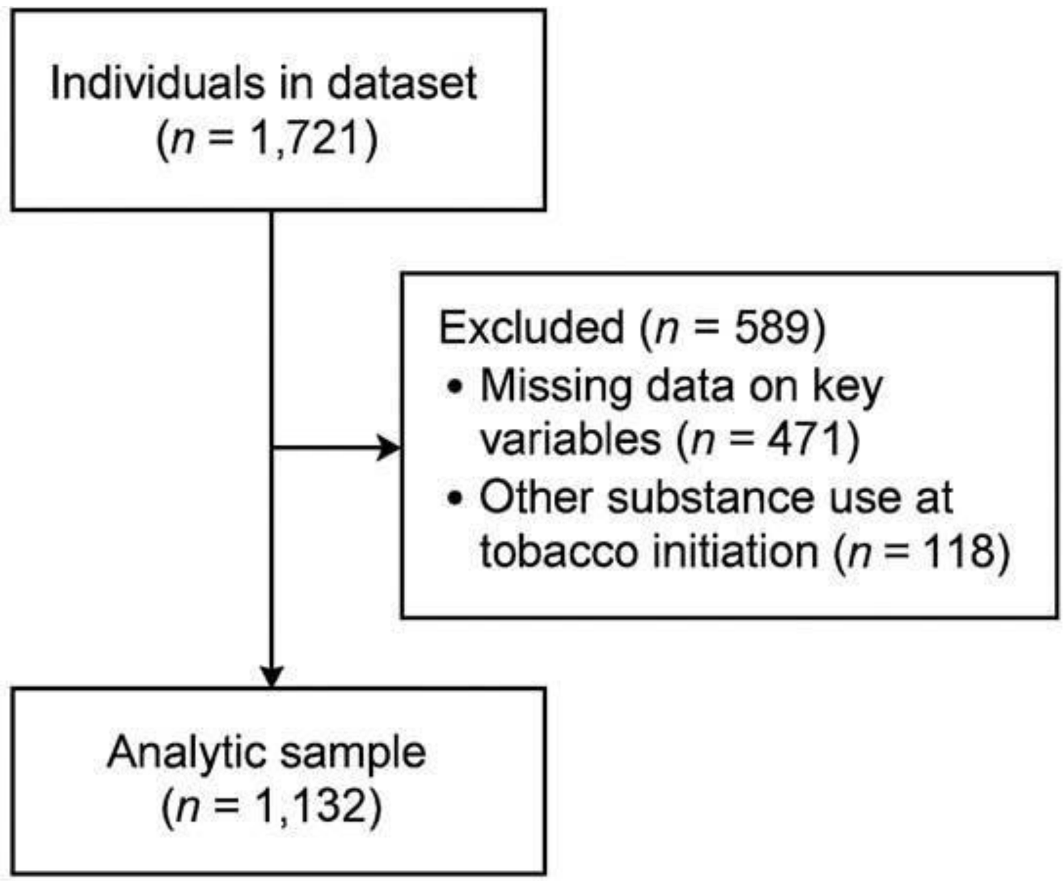


Figure 1. Flowchart illustrating inclusion and exclusion criteria and final analytic sample.

Data were collected on two primary outcomes: age of tobacco initiation and mental health symptoms. Tobacco use initiation was measured through self-reported age at first use and categorized into early initiation (17 years or younger) and later initiation (older than 17). Anxiety symptoms were assessed using the Generalized Anxiety Disorder-7 (GAD-7; Spitzer et al., 2006), while depressive symptoms were measured using the Patient Health Questionnaire-9

(PHQ-9; Kroenke et al., 2001). Both screening tools have been widely validated in general and adolescent populations and are commonly used in epidemiological research to assess mental health symptoms. Their inclusion allows for the identification of moderate to severe symptoms without requiring formal clinical diagnoses (Spitzer et al., 2006).

Additional demographic information was collected, including participants' age, sex, socioeconomic status, and educational levels, as these factors may contribute to anxiety and depression. These variables were included to explore potential confounders, such as whether socioeconomic status or sex alters the relationship between tobacco use and mental health symptoms.

Chi-square tests were conducted to compare the prevalence of anxiety and depression between early and later tobacco initiators. Separate chi-square tests were also conducted within sex groups to determine whether these relationships differed by sex. These analyses were designed to assess both the overall and sex-specific associations between tobacco initiation and mental health outcomes.

All statistical analyses were conducted using IBM SPSS Statistics (version 29). Descriptive statistics were used to summarize participant characteristics, with categorical variables reported as frequencies and percentages. Chi-square tests were used for categorical comparisons across groups, and statistical significance was set at $p < 0.05$.

Tests for normality for the two primary outcomes were not required, as these variables were categorical. Missing data were handled using listwise deletion. Participants with missing values on any key variables (such as tobacco use, anxiety or depression score, or sex) were excluded prior to analysis, as illustrated in **Figure 1**.

Results

A total of 1,132 participants were included in the final analysis. Participants were categorized as early initiators (≤17 years) or late initiators (>17 years) and stratified by sex. Out of the original 1,721 participants, 471 individuals were excluded due to missing data on tobacco use initiation, mental health symptoms, or sex. An additional 118 participants were excluded for reporting concurrent use of other substances at the time of tobacco initiation. **Table 1** presents descriptive characteristics, including demographic, smoking behaviour, and mental health-related variables.

Table 1.

Baseline characteristics of participants of the 2022 Mental Health and Access to Care Survey, categorized by sex and age of tobacco initiation.

Variable	Early Initiation (≤17 years) (n=571)	Late Initiation (>17 years) (n=561)
Sex presented as *n (%)*		
Male	285 (49.9%)	287 (51.2%)
Average age of different smoking habits (years)		

Smoked first whole cigarette[a]	14.3 (3)	18.8 (2)
Started smoking daily[a]	15.8 (3)	19.3 (4)
Started smoking daily - Former daily smoker[b]	17.9 (2.5)	22.3 (2.8)
Educational level completed presented as *n (%)*		
No High School	52 (9.1%)	40 (7.1%)
High School	181 (31.7%)	179 (31.9%)
Some College/University	218 (38.2%)	224 (39.9%)
College/University Graduate	120 (21.0%)	118 (21.0%)
Household income level presented as *n (%)*		

Low	182 (31.9%)	171 (30.5%)
Middle	238 (41.7%)	238 (42.4%)
High	151 (26.4%)	152 (27.1%)
Categories of smoking habits presented as *n (%)*		
Smoked 100+ Cigarettes (Yes)	472 (82.7%)	511 (91.1%)
Type of Smoker - Daily	197 (34.5%)	227 (40.5%)
Type of Smoker - Occasional	163 (28.5%)	165 (29.4%)
Type of Smoker - Former	69 (12.1%)	65 (11.6%)
Mental health indicators presented as *n (%)*		

Depression (Yes)	374 (65.5%)	357 (63.6%)
Anxiety (Yes)	320 (56.0%)	329 (58.6%)

Description of sex, mental health indicators, smoking habits, educational levels and household income among early (≤17 years) and late (≥18 years) initiators of tobacco usage. Values for categorical variables are represented through counts and percentages, while values for continuous variables are represented through measures of central tendency, determined from normality tests.

[a] Measure of central tendency reported for this continuous variable is Median (IQR).

[b] Measure of central tendency reported for this continuous variable is Mean (SD).

As shown in the table, participants were relatively evenly distributed across initiation groups and sex. The majority had completed at least some form of education beyond high school, and most reported moderate-income levels. Daily smoking was reported more frequently among late initiators. The average age of smoking initiation among early initiators was around 14 years, while the average age among late initiators was around 19 years. Notably, over 80% of the total sample reported smoking more than 100 cigarettes in their lifetime.

Figures 2 and 3 visually highlight the distribution of anxiety and depression symptoms across subgroups.

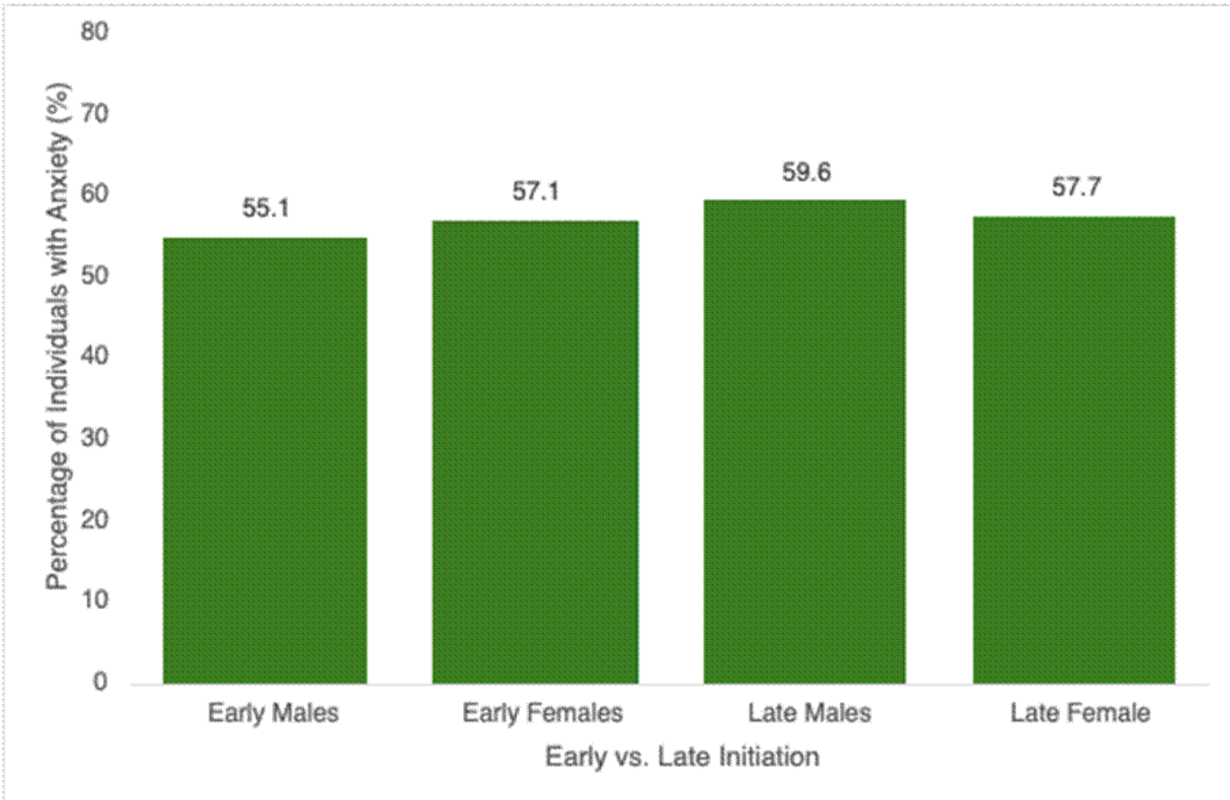


Figure 2. Percentage of individuals reporting symptoms of anxiety (GAD-7) by tobacco initiation group and sex.

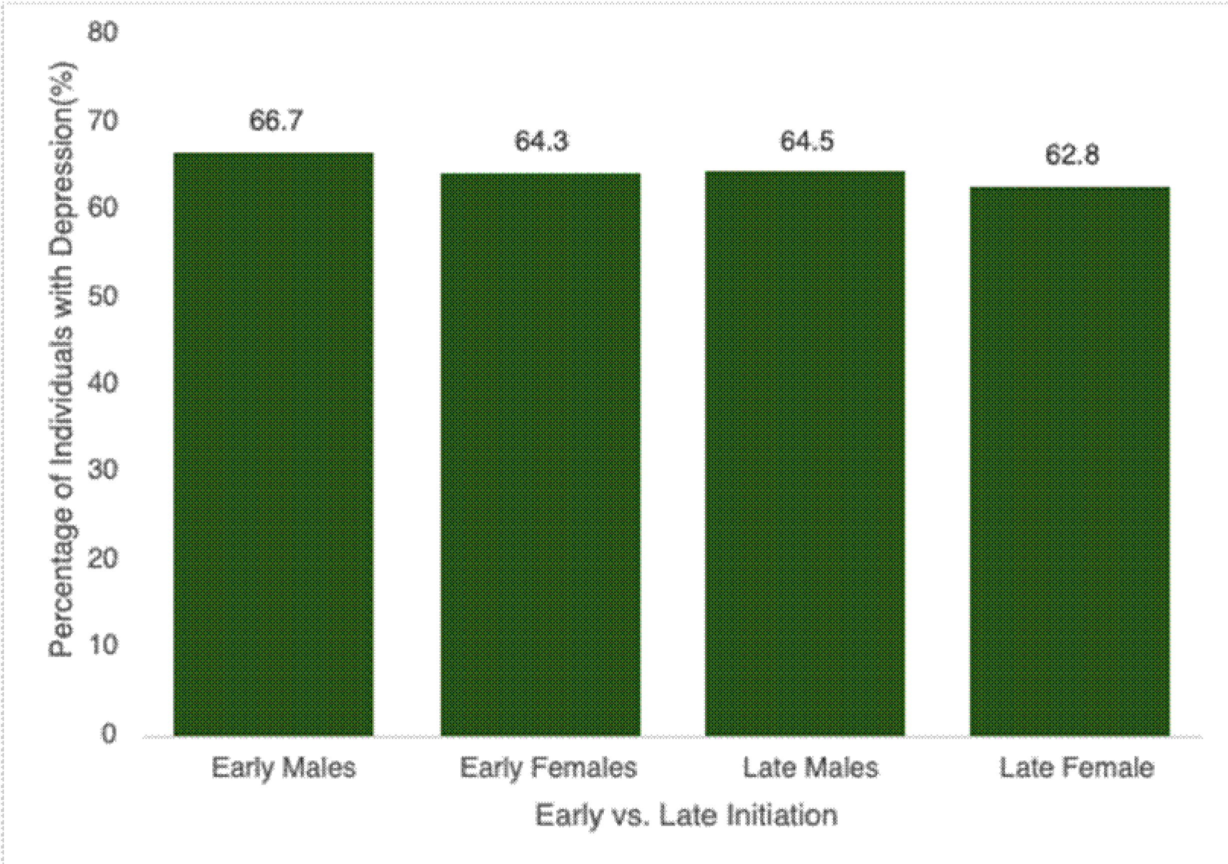


Figure 3. Percentage of individuals reporting symptoms of depression (PHQ-9) by tobacco initiation group and sex.

The chi-square analysis for anxiety revealed no significant association between tobacco initiation group and anxiety symptoms ($\chi^2(3, N = 1132) = 1.21, p = .751$). **Figure 2** shows that anxiety prevalence ranged from roughly 55% to 60% across all groups, with late-initiating males reporting slightly higher symptoms than other groups.

Similarly, the analysis for depression showed no statistically significant association ($\chi^2(3, N = 1132) = 0.94, p = .815$). As shown in **Figure 3**, reported depression symptoms ranged roughly between 63% and 67%, with early-initiating males showing slightly higher rates than other subgroups.

No significant interaction was found between sex and initiation age for either anxiety or depression. Specifically, the sex-stratified chi-square test for depression resulted in $\chi^2(1, N = 1132) = 0.42$, $p = .525$, and for anxiety, $\chi^2(1, N = 1132) = 0.12$, $p = .978$. These findings suggest that differences in mental health symptoms across groups were small and not statistically significant.

Significant findings would have been indicated with an asterisk (*) in the figures. However, as there were no significant findings, no such notation was needed.

Discussion

The primary goal of this study was to investigate whether individuals who initiated tobacco use at age 17 or younger had a higher prevalence of self-reported depression and anxiety compared to those who began later. The statistical analysis showed no statistically significant differences in the prevalence of depression or anxiety between early and late initiators of tobacco. Furthermore, there was no significant difference found between sexes in reported mental health symptoms across both groups. These findings suggest that age of tobacco initiation, on its own, may not be a strong independent predictor of depression or anxiety.

Previously published studies have shown that tobacco use can significantly increase susceptibility to mood disorders (Benowitz, 2010). In addition, smoking cigarettes has also been associated with the reduction of several brain structures, including the prefrontal cortex, which is the last part of the brain to fully develop (Elbejjani et al., 2019). Based on these findings, we hypothesized that individuals who began smoking at an earlier age (during adolescence, when brain development is still ongoing) would show higher rates of anxiety and depression. This expectation was also based on research suggesting that early nicotine exposure may increase vulnerability to stress-related disorders (Rohleder & Kirschbaum, 2006). Given that adolescence is a critical window for neurodevelopment, especially in areas involved in emotion and executive functioning, early tobacco use was expected to carry greater mental health risks compared to later initiation. While our primary focus was on initiation age, we also examined whether these relationships differed by sex. We hypothesized that females would report higher symptom levels, as they tend to show higher rates of depression and anxiety during adolescence and may be more sensitive to nicotine's effects on stress and mood regulation (Jandíková et al., 2017).

However, the results did not align with previous literature. While there were small differences in the proportions of individuals experiencing symptoms, they were not statistically significant. This discrepancy may stem from methodological differences. The studies we are comparing to used longitudinal designs and clinical diagnoses, while ours was cross-sectional and relied on self-reported screening tools. This means that our data cannot be used to establish causality, and it may not have captured more subtle variations of mental health symptoms compared to studies using clinical assessments (made by psychologists and psychiatrists). Also, the sample type used in our study differed from the adolescent high-risk groups that were studied in the papers we examined. Using a general population sample from a national survey may help explain the lack of statistically significant findings. It is also worth noting that studies based only on high-risk populations may not be representative of the general public, which presents a limitation in their applicability.

Despite these differences, a key strength of our study is the dataset: the 2022 Mental Health and Access to Care Survey. It includes a large, nationally representative sample, which increases the generalizability and external validity of the findings. It also uses validated screeners like the PHQ-9 and GAD-7, which have been shown to have high consistency and validity across diverse populations. Lastly, it captures data on age of initiation, frequency of smoking, and long-term patterns, allowing for nuanced subgroup analyses. This degree of detail in tobacco use history is often missing in studies that only capture current smoking status.

As with all studies, there were limitations. First, a cross-sectional design cannot be used to infer causality. This means that we do not know whether tobacco use contributed to mental health symptoms, or whether those symptoms existed prior to smoking initiation. Second, our reliance on self-reported data introduces potential for recall bias and social desirability bias. In this study in particular, social desirability bias (defined as a tendency to report information in a way that aligns with perceived social norms) may have influenced participants' responses, particularly due to stigma around smoking and mental health (Bergen & Labonté, 2020). Participants may have underreported tobacco use or mental health symptoms due to perceived stigma or discomfort discussing these topics. Third, we used a binary grouping for initiation age. This may oversimplify a more nuanced relationship between tobacco initiation age and mental health symptoms. For instance, someone who started smoking at 13 years old may have developmental or contextual differences compared to someone who started smoking at 17 years old, even though they would be grouped together in the analysis.

To address these limitations, future researchers should use longitudinal data, such as the National Longitudinal Survey of Children and Youth, to assess how smoking onset and mental health symptoms evolve over time. This would allow researchers to assess the temporal sequence of associations between tobacco use and mental health symptoms. Furthermore, using time-to-event analyses like Cox regression can help clarify when mental health symptoms emerge in relation to smoking (Zhang, 2018). It is also important to incorporate biomarkers to validate smoking status. Biomarkers (like nicotine metabolites) provide objective physiological evidence of tobacco exposure and reduce reliance on self-reported data, which can be prone to recall bias. If self-reported data are used, however, researchers should use anonymous or digital survey tools to reduce social desirability bias. Finally, future studies should treat initiation age as a continuous variable, and also factor in frequency and intensity of use to more accurately capture patterns of tobacco exposure and its associations with mental health outcomes.

Our findings suggest that the link between tobacco initiation age and mental health may be more complex than originally thought, and not solely driven by age or sex. This suggests that public health strategies aimed at reducing tobacco-related health risks may benefit from addressing a broader range of factors (like frequency of use, co-occurring stressors, and social context) rather than focusing solely on the age of initiation.

References

Bergen, N., & Labonté, R. (2020). "Everything Is Perfect, and We Have No Problems": Detecting and Limiting Social Desirability Bias in Qualitative Research. Qualitative health research, 30(5), 783–792. https://doi.org/10.1177/1049732319889354

Bruijnzeel A. W. (2012). Tobacco addiction and the dysregulation of brain stress systems. *Neuroscience and biobehavioral reviews*, *36*(5), 1418–1441. https://doi.org/10.1016/j.neubiorev.2012.02.015

Dai, H. D., Doucet, G. E., Wang, Y., Puga, T., Samson, K., Xiao, P., & Khan, A. S. (2022). Longitudinal Assessments of Neurocognitive Performance and Brain Structure Associated With Initiation of Tobacco Use in Children, 2016 to 2021. *JAMA Network Open*, *5*(8), e2225991–e2225991. https://doi.org/10.1001/jamanetworkopen.2022.25991

Elbejjani, M., Auer, R., Jacobs, D. R., Haight, T., Davatzikos, C., Goff, D. C.,Bryan, R. N., & Launer, L. J. (2019). Cigarette smoking and gray matter brain volumes in middle age adults: the CARDIA Brain MRI sub-study. *Translational Psychiatry*, *9*(1). https://doi.org/10.1038/s41398-019-0401-1

Goriounova, N. A., & Mansvelder, H. D. (2012). Short- and long-term consequences of nicotine exposure during adolescence for prefrontal cortex neuronal network function. *Cold Spring Harbor perspectives in medicine*, *2*(12), a012120. https://doi.org/10.1101/cshperspect.a012120

Hajdusianek, W., Żórawik, A., Waliszewska-Prosół, M., Poręba, R., & Gać, P. (2021). Tobacco and Nervous System Development and Function—New Findings 2015–2020. *Brain Sciences*, *11*(6), 797. https://doi.org/10.3390/brainsci11060797

Jandíková, H., Dušková, M., & Stárka, L. (2017). The influence of smoking and cessation on the human reproductive hormonal balance. Physiological research, 66(Suppl 3), S323–S331. https://doi.org/10.33549/physiolres.933724

Kroenke, K., Spitzer, R.L. & Williams, J.B.W. The PHQ-9. *J GEN INTERN MED* 16, 606–613 (2001). https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Laviolette, S. R. (2021). Molecular and neuronal mechanisms underlying the effects of adolescent nicotine exposure on anxiety and mood disorders. *Neuropharmacology, 184,* 108411. https://doi.org/10.1016/j.neuropharm.2020.10841

Lutfy, K., Aimiuwu, O., Mangubat, M., Shin, C.-S., Nerio, N., Gomez, R., Liu, Y., & Friedman, T. C. (2012). Nicotine stimulates secretion of corticosterone via both CRH and AVP receptors. *Journal of Neurochemistry*, no-no. https://doi.org/10.1111/j.1471-4159.2011.07633.

Fluharty, Amy E. Taylor, Meryem Grabski, Marcus R. Munafò, (2017)*The Association of Cigarette Smoking With Depression and Anxiety: A Systematic Review, Nicotine & Tobacco Research, Volume 19, Issue 1, 1 January 2017, Pages 3–13, https://doi.org/10.1093/ntr/ntw140*

Mental Health Foundation. (2023). Smoking and mental health. Mental Health Foundation; MHF. https://www.mentalhealth.org.uk/explore-mental-health/a-z-topics/smoking-and-mental-health?utm_source=chatgpt.com

Rohleder, N., & Kirschbaum, C. (2006). *The hypothalamic-pituitary-adrenal (HPA) axis in habitual smokers. International journal of psychophysiology : official journal of the International Organization of Psychophysiology, 59(3), 236–243. https://doi.org/10.1016/j.ijpsycho.2005.10.012*

Rosa Alati, Stuart Kinner, Jake M. Najman, Greg Fowler, Kerrianne Watt, David Green, GENDER DIFFERENCES IN THE RELATIONSHIPS BETWEEN ALCOHOL, TOBACCO AND MENTAL HEALTH IN PATIENTS ATTENDING AN EMERGENCY DEPARTMENT, *Alcohol and Alcoholism*, Volume 39, Issue 5, September/October 2004, Pages 463–469, https://doi.org/10.1093/alcalc/agh080

Shi, P., Yang, A., Zhao, Q., Chen, Z., Ren, X., & Dai, Q. (2021). A Hypothesis of Gender Differences in Self-Reporting Symptom of Depression: Implications to Solve Under-Diagnosis and Under-Treatment of Depression in Males. *Frontiers in Psychiatry*, *12*(12). https://doi.org/10.3389/fpsyt.2021.589687

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. Archives of Internal Medicine, 166(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Stephenson, E. (2023, September 22). Mental disorders and access to mental health care. https://www150.statcan.gc.ca/n1/pub/75-006-x/2023001/article/00011-eng.htm

Thompson, A. B., Tebes, J. K., & McKee, S. A. (2016). Gender differences in age of smoking initiation and its association with health. *Addiction Research & Theory*, *23*(5), 413–420. https://doi.org/10.3109/16066359.2015.1022159

Valentine, G., & Sofuoglu, M. (2018). Cognitive Effects of Nicotine: Recent Progress. *Current Neuropharmacology*, *16*(4), 403–414. https://doi.org/10.2174/1570159x15666171103152136

World Health Organization. (2024, October 10). *Mental Health of Adolescents*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health

Wu, A. D., Gao, M., Aveyard, P., & Taylor, G. (2023). Smoking Cessation and Changes in Anxiety and Depression in Adults With and Without Psychiatric Disorders. *JAMA Network Open*, *6*(5), e2316111. https://doi.org/10.1001/jamanetworkopen.2023.16111

Zhang, Z., Reinikainen, J., Adeleke, K. A., Pieterse, M. E., & Groothuis-Oudshoorn, C. G. M. (2018). Time-varying covariates and coefficients in Cox regression models. Annals of translational medicine, 6(7), 121. https://doi.org/10.21037/atm.2018.02.12

Exploring the Relationship Between Osteoporosis and Recreational Activity Among Canadian Adults

Rubab Masaba, Andrew Son, Nikou Zahedi

McMaster University

LIFESCI 3LX3: Applied Biostatistics & Scientific Writing

Dr. Janet Pritchard

April 08, 2025

Introduction

Osteoporosis is a progressive skeletal disorder characterized by decreased bone mineral density (BMD) and deterioration of bone microarchitecture. This condition leads to increased bone fragility and a heightened risk of fractures. In Canada, about 1.5 million individuals over the age of 40 nearly 10% of this age group live with a diagnosis of osteoporosis (McArthur et al., 2022). Women face this diagnosis at nearly four times the rate of men, revealing a significant sex disparity that places them at disproportionately greater risk. As Canada's population continues to age, the public health burden of osteoporosis will likely grow in scope and impact. This demographic trend demands urgent attention to the lived experiences, behaviors, and health outcomes of individuals managing this condition on a daily basis.

Healthcare professionals consistently recommend regular physical activity, particularly weight-bearing and resistance-based exercises, as a primary non-pharmacological strategy for osteoporosis management (Howe et al., 2011). These exercises play a critical role in preserving bone strength, maintaining mobility, and reducing the risk of fractures. However, data show that fewer than half of older Canadian adults meet basic physical activity guidelines (Howe et al., 2011). Research has shown that individuals with osteoporosis tend to experience diminished physical fitness, decreased muscular strength, impaired balance, and heightened fear of falling (Babiuch et al., 2021; Meyer et al., 2019). These issues do not operate in isolation, they combine to create complex physical and psychological barriers to activity. Babiuch et al. (2021) found that women between 60 and 65 years old with osteoporosis scored significantly lower on standard physical fitness assessments and appeared at greater risk for serious falls compared to their peers without the condition. Meyer et al. (2019) observed that even individuals without a history of falls often restrict their movement due to the psychological fear of falling. This fear frequently leads to avoidance behaviors, resulting in deconditioning, declining mobility, and further withdrawal from physical activity. Such patterns reflect a self-reinforcing cycle of inactivity that may exacerbate existing health risks and limit functional independence.

Despite these obstacles, many people with osteoporosis do remain physically active by adapting their exercise choices. They often gravitate toward low-impact, non-competitive, and home-based activities that they perceive as safer and more manageable. Kelley (2018) reported that walking represents the most common form of recreational activity among older adults with osteoporosis. Gardening, yard work, and self-guided home-based routines closely follow. In contrast, very few individuals in this group engage in higher-impact or more physically demanding activities, such as jogging, team sports, or aerobic classes (Florvåg et al., 2025). Florvåg et al. (2025) emphasized that people generally select activities that align with their comfort, confidence, and perceived safety. These preferences lead individuals to stay active within a narrow range of movement, often excluding activities that involve risk, complexity, or unfamiliar environments.

Physical impairments linked with osteoporosis further shape activity preferences. Progressive muscle weakness, spinal kyphosis, and chronic balance challenges limit functional mobility, especially in aging women who face an elevated risk of vertebral and nonvertebral fractures (Babiuch et al., 2021). These impairments also affect self-confidence, potentially contributing to reduced participation in group-based or outdoor exercise settings. However, consistent evidence shows that individuals who maintain regular physical activity experience notable improvements in both function and autonomy (Meyer et al., 2019). Exercise supports postural stability, coordination, and neuromuscular control all of which are critical for preventing falls and preserving daily independence.

Structured exercise interventions have repeatedly demonstrated their effectiveness in improving health outcomes in osteoporotic populations. Stanghelle et al. (2020) found that a 12-week resistance and balance program significantly improved muscular strength, mobility, and fear-of-falling scores among older women with vertebral fractures. In addition to enhancing BMD, structured programs promote weight regulation, improve cardiovascular endurance, reduce depressive symptoms, and increase opportunities for social interaction (Babiuch et al., 2021). These benefits extend far beyond skeletal health, emphasizing the value of movement for whole-body well-being.

Meta-analytic evidence further confirms the physiological effects of physical activity on bone outcomes. Su et al. (2020) evaluated the impact of swimming on lumbar spine BMD and observed significantly higher BMD levels among swimmers than non-swimmers. Postmenopausal women who swam for three to six hours weekly exhibited especially strong gains. Su et al. (2020) noted that these effects were not apparent in premenopausal swimmers or in those with shorter swim durations, highlighting that the impact of exercise varies with biological and behavioral factors such as age, sex, and intensity. Although swimming is not a weight-bearing activity, this finding reinforces the value of consistent movement and stratified behavioral analysis when evaluating intervention effects in osteoporotic populations.

Nonetheless, despite the wealth of evidence supporting physical activity, many individuals with osteoporosis remain underactive. Several well-documented barriers contribute to this trend. Pain, comorbid health conditions, and confusion about what constitutes “safe” exercise often deter participation (Meyer et al., 2019; Simmonds et al., 2015). Simmonds et al. (2015) also reported that patients frequently receive inconsistent or unclear guidance from healthcare providers, which can lower their motivation and increase anxiety about movement.

Environmental and logistical challenges such as lack of transportation, cost of programming, and absence of age-friendly fitness spaces further complicate participation for many older adults, particularly those in rural or underserved communities.

Although interest in physical activity as a prevention and treatment strategy for osteoporosis has grown, the literature remains limited in several key areas. Most studies continue to emphasize pre-diagnosis behaviors or examine exercise in controlled clinical contexts, often overlooking how individuals adjust their behavior following diagnosis (Nguyen, 2016). Moreover, while researchers frequently discuss sex-based differences in osteoporosis prevalence and psychosocial responses such as fear of falling, they rarely evaluate how these differences translate into measurable differences in recreational activity levels (Aoyagi & Shephard, 2013). Similarly, few population-based studies examine how people with osteoporosis perceive and navigate social or environmental constraints that limit their participation. Without this knowledge, efforts to promote activity may fail to address the real barriers that individuals experience, especially among women and other high-risk subgroups.

To address these gaps, the present study uses data from the 2015 Canadian Community Health Survey to examine the relationship between osteoporosis and self-reported recreational activity patterns in a nationally representative adult population. The first research question asks: Among Canadian adults aged 40 and older, does having a concurrent diagnosis of osteoporosis affect the number of minutes spent on recreational activities per week? This question explores whether osteoporosis status correlates with reduced participation in leisure-time physical activity. The second question examines sex-based differences: Among Canadian adults aged 40 and older with a concurrent diagnosis of osteoporosis, do females, compared to males, report lower activity levels and spend fewer minutes on recreational activities per week?

Answering these questions can provide critical insights into how osteoporosis affects day-to-day activity behavior and how these patterns differ across sex groups. This information can guide the development of tailored public health strategies, more inclusive physical activity guidelines, and accessible community-based programs. If the data confirm that women with osteoporosis participate less in recreational activity, public health officials, clinicians, and program developers can design interventions that directly address the unique challenges this population faces. By supporting sustained, safe, and enjoyable physical activity, we can help individuals with osteoporosis maintain functional independence, reduce fall risk, and enhance their overall quality of life.

Methods

Our study utilized a cross-sectional design, based on secondary analysis of publicly available data from the 2015 Canadian Community Health Survey (CCHS) – accessed through ODESI. The CCHS– is a nationally representative survey conducted by Statistics Canada to assess Canadians' food and nutrient intake, as well as general health and lifestyle behaviors. Although the primary focus of the dataset is on dietary intake, our study specifically used data from the general health questionnaire, which gathers information on chronic conditions (such as osteoporosis), physical activity patterns, and sociodemographic characteristics. The survey employed a multi-stage, stratified cluster sampling design and included both in-person and telephone interviews. To ensure national representativeness, Statistics Canada provided survey weights, and all data were collected under the authority of the Statistics Act of Canada.

For the study population, we used a certain criteria to determine what would be included and excluded from our data. For instance, we decided to focus on respondents only from Alberta, as the sample size from Nunavut was insufficient for meaningful analysis. Thus, we excluded participants that were from Nunavut from our data. In addition, we only included participants aged 40 years and older, given their increased risk for osteoporosis, aligning with our research objectives. The last piece of criteria we used for inclusion and exclusion was that we excluded individuals if they had reported implausible or extreme outliers in physical activity data (e.g., over 2000 minutes per week) or if they had missing responses for osteoporosis diagnosis or recreational activity. In result, we excluded individuals from our data if they resided in Nunavut, if they were under 40 years old, or if they reported abnormal data. The final sample for Research Question 1 included 6,870 participants, while the subsample for Research Question 2, focused on individuals with a self-reported osteoporosis diagnosis, included 640 participants.

The primary outcome variable for this study was osteoporosis diagnosis, which was self-reported by participants in response to the question, "Do you have osteoporosis?" Responses were coded as “Yes” (1) for osteoporosis and “No” (0) for no osteoporosis. This self-reported measure has shown reasonable validity when compared to clinical records, particularly in large epidemiological studies. The primary independent variable was recreational physical activity, assessed through the question, "In the past 7 days, how many minutes did you spend doing leisure-time physical activities (e.g., walking, gardening, cycling, sports)?" Responses were given as a continuous variable in minutes per week and were then categorized into three groups: 0 minutes/week, 1–359 minutes/week, and ≥360 minutes/week. The ≥360 minutes/week threshold aligns with the Canadian Physical Activity Guidelines, which recommends at least 150 minutes of moderate-to-vigorous activity per week.

To account for confounding variables and provide a more nuanced interpretation of the data, we included several socio demographic covariates: age (both continuous and categorical, e.g., 40–44, 45–49, etc.), sex (male/female), ethnicity (White/Non-White), and marital status (Married/Common-law, Single, Divorced, Widowed, Separated). We also included non-demographic covariates such as denture usage and food insecurity. These factors have shown to influence both osteoporosis risk and physical activity patterns.

For statistical analysis, all procedures were conducted using IBM SPSS Statistics, with a two-tailed p-value < 0.05 considered statistically significant. Data cleaning and coding were completed before analysis, and missing values were excluded using listwise deletion. Categorical variables (such as sex and osteoporosis status) were summarized with frequencies and percentages. For continuous variables (like minutes of activity), we tested for normality using histograms, Q-Q plots, and the Kolmogorov-Smirnov test, which indicated that recreational activity data were not normally distributed. Consequently, we employed non-parametric tests.

For bivariate analysis, we used chi-square tests to examine relationships between categorical variables, such as osteoporosis status vs. activity level ($<$360 vs. $\geq$360 minutes per week), and gender vs. activity level among individuals with osteoporosis. We used independent-sample t-tests or Mann-Whitney U tests for continuous comparisons, such as comparing age across groups. For multivariable analysis, we applied a binary logistic regression model to assess the relationship between recreational activity level and the presence of osteoporosis, adjusting for potential confounders (including age, sex, income, education, marital status, etc.). We reported results as adjusted odds ratios (AORs) with 95% confidence intervals. To ensure the validity of the model, we assessed multicollinearity using variance inflation factors (VIFs).

Results

We began by examining whether osteoporosis status was associated with recreational physical activity in the full sample (N = 6870). As shown in (Figure.1), participants who engaged in less than 360 minutes of activity per week had a slightly higher prevalence of osteoporosis (9.49%) than those exceeding that threshold (7.00%). However, this difference was not statistically significant ($\chi^2(1) = 3.04$, $p = .081$).

To assess broader predictors of high activity levels (≥360 min/week), we conducted a logistic regression analysis in the full sample. The model was statistically significant ($\chi^2(10) = 63.28$, $p < .001$). The Hosmer–Lemeshow test indicated good model fit ($\chi^2(7) = 9.542$, $p = .216$), suggesting that the model adequately represented the data. Age group, denture use, and cultural background significantly predicted activity level. Adults aged 70 and over had lower odds of meeting the 360-minute threshold compared to those aged 40-55. Participants who did not wear dentures and those identifying as White were more likely to be highly active. Osteoporosis diagnosis did not significantly predict activity. Full results are presented in (Supplementary Table 2).

We further explored sex differences in osteoporosis diagnosis. As shown in (Figure 2), women made up 84.2% of participants diagnosed with osteoporosis, while only 15.8% were men. A chi-square test confirmed that this difference was statistically significant ($\chi^2(1) = 276.15$, $p < .001$), indicating a higher prevalence of osteoporosis among women as compared to men.

Next, we examined whether sex was associated with recreational activity levels within the osteoporosis subsample (N = 640). (Figure 3) shows that 9.9% of males with osteoporosis reported engaging in more than 360 minutes of recreational activity per week, compared to only

4.6% of females. This difference was statistically significant ($\chi^2(1) = 4.56$, $p = .033$), suggesting that osteoporotic males were more likely to maintain high levels of physical activity than osteoporotic females.

Finally, we conducted a logistic regression restricted to participants with osteoporosis to identify predictors of high activity in the group. The model was statistically significant ($\chi^2(11) = 78.34$, $p < .001$). The Hosmer–Lemeshow test indicated excellent fit ($\chi^2(8) = 0.270$, $p = 1.000$), supporting the model's reliability in this subsample. Age group was the only significant predictor: individuals aged 70 and over were significantly less likely to meet the recommended activity threshold compared to those aged 40-55 ($p = .036$). Other variables, including sex, denture use, food insecurity, and cultural background, were not significant. Detailed results are available in (Supplementary Table 3).

Table 1. Participant Characteristics by Osteoporosis Diagnosis (N = 6870)

Descriptive distribution of demographic, behavioral, and health variables stratified by osteoporosis status. Values represent the percentage of participants within each category. Categories include age group, sex, recreational activity level, denture use, cultural/racial background, marital status, food insecurity, and alternate physical activity classification.

	Categories	Osteoporosis	No Osteoporosis
Age	40-55	0.8%	37.4%
	55-70	4.3%	35.5%
	>70	4.3%	17.8%
Sex	Male	1.5%	45.5%
	Female	7.8%	45.2%
Recreational Activity	0 min/week	6.2%	54.1%
	0-<360 min/week	2.6%	29.9%
	>360 min/week	0.5%	6.7%
Denture	Yes	3.9%	23.2%
	No	5.4%	67.5%
Culture/Racial Background	White	8.6%	81.0%
	Non-White	0.7%	9.7%
Marital Status	Married	4.2%	52.2%
	Common-Law	0.3%	5.5%
	Widowed/Divorced/Separated	4.1%	22.7%
	Single	0.7%	10.3%
Food Insecurity	Always	0.2%	1.1%
	Sometimes	0.5%	4.2%
	Never	8.6%	85.3%
Alternate physical activity indicator	Active	4.0%	49.7%
	Somewhat Active	2.5%	20.8%
	Sedentary	2.8%	20.2%

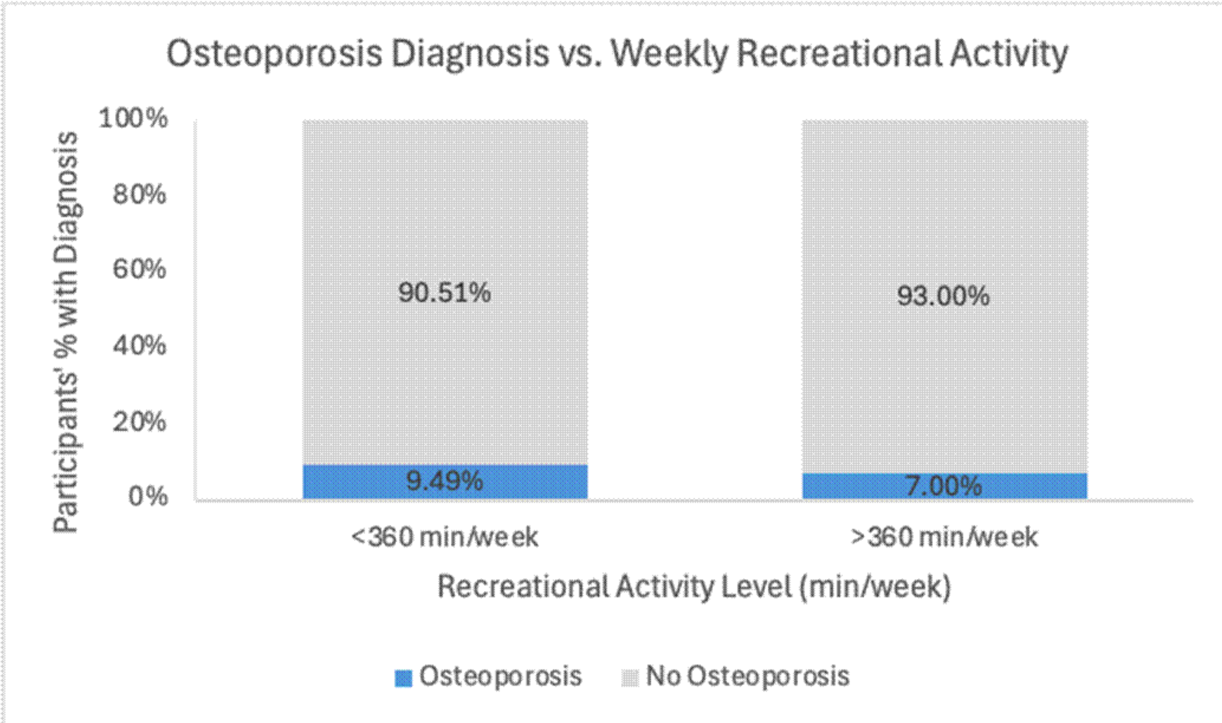


Figure 1. Osteoporosis Prevalence by Recreational Activity Level (N = 6870)

Proportion of participants with and without osteoporosis by weekly recreational activity level. Among the total sample (N = 6870), those engaging in less than 360 minutes of activity per week had slightly higher prevalence of osteoporosis (9.49%) compared to those exceeding 360 minutes (7.00%). This difference was not statistically significant ($\chi^2(1) = 3.04$, $p = .081$).

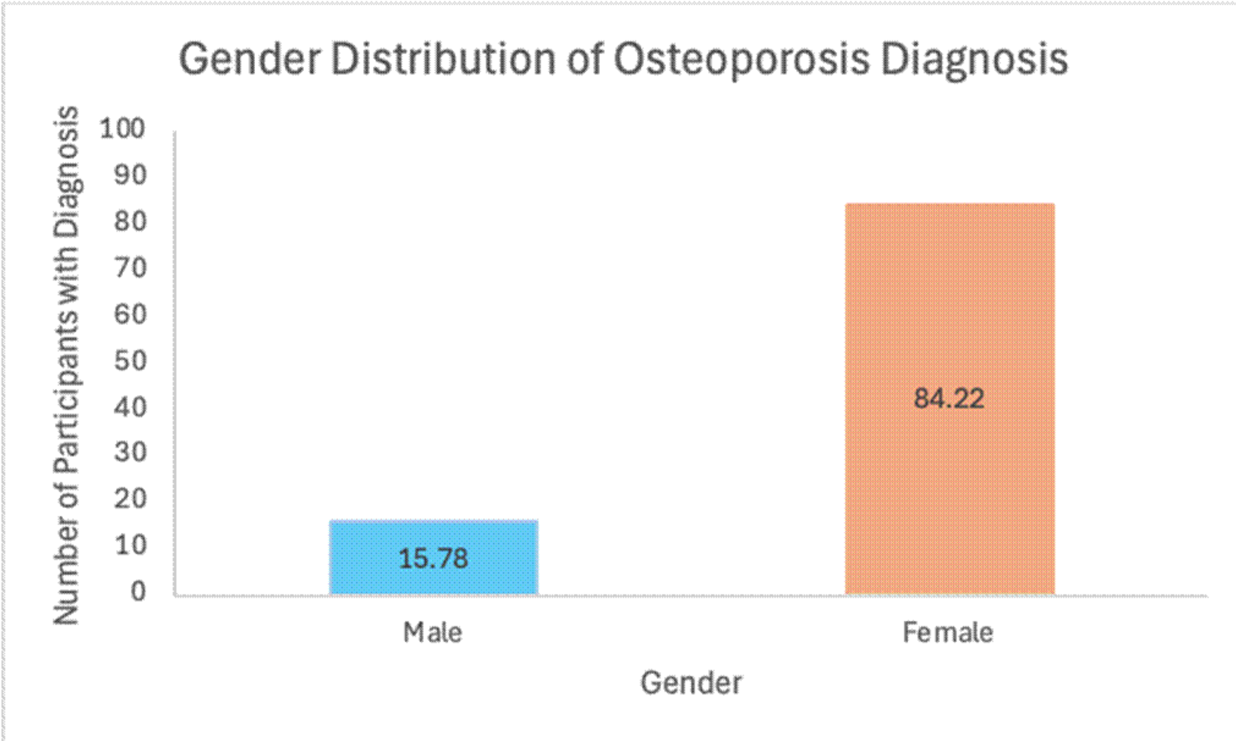


Figure 2. Gender Distribution of Osteoporosis Diagnosis (N = 640)

Among participants diagnosed with osteoporosis (N = 640), the majority were female consisting of 84.22%, while only 15.78% were male. A chi-square test confirmed a significant association between gender and osteoporosis diagnosis, indicating a higher prevalence among women ($\chi^2(1) = 276.15$, $p < .001$).

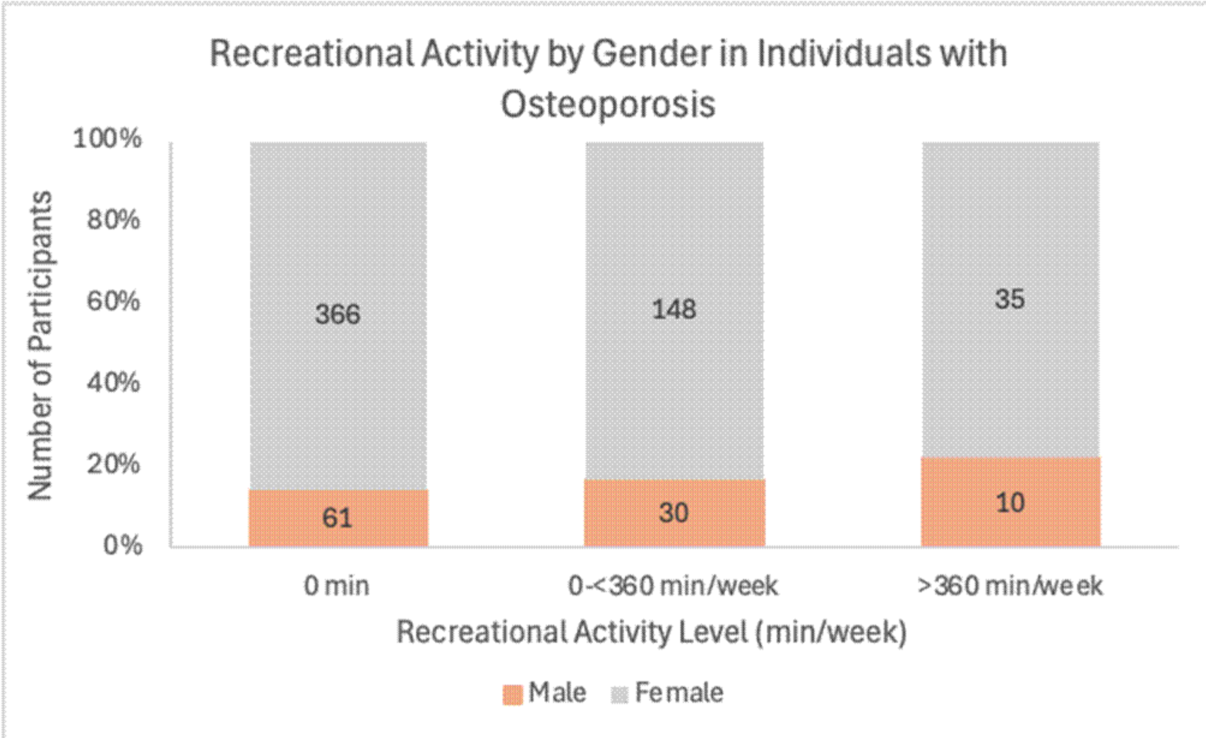


Figure 3. Recreational Activity by Gender in Individuals with Osteoporosis (N = 640)

Stacked bar chart showing the distribution of recreational activity levels among males and females diagnosed with osteoporosis. A higher proportion of males reported more than 360 minutes of weekly activity compared to females (9.9% vs. 4.6%). A chi-square test indicated a significant association between gender and recreational activity level in this subsample ($\chi^2(1)$ = 4.56, p = .033).

Discussion

Our analysis of predictors of recreational physical activity revealed several important insights. In the full sample of 6,870 participants, we used a logistic regression model to identify factors associated with engaging in high levels of physical activity, defined as 360 minutes or more per week. This model turned out to be statistically significant and demonstrated good fit, indicating that it accurately captured relationships in the data. Three variables emerged as significant predictors of high activity levels: age, denture use, and cultural background. Specifically, adults aged 70 and over were significantly less likely to meet the recommended activity threshold compared to those aged 40 to 55. Individuals who did not wear dentures were more likely to be physically active, potentially reflecting better overall health. Additionally, participants who identified as White were more likely to engage in high levels of activity than those from other cultural backgrounds. These findings suggest that age, denture use, and ethnicity are valid and meaningful predictors of physical activity in the general population.

Contrary to expectations, osteoporosis status was not a significant predictor of physical activity. Although there was a slightly lower prevalence of osteoporosis among those who were more physically active (7.00%) compared to those who were less active (9.49%), this difference was not statistically significant. When included in the logistic regression model, osteoporosis diagnosis did not significantly influence whether someone met the 360-minute activity threshold. This indicates that, after controlling for other variables, having osteoporosis alone does not reliably predict activity levels.

Further analyses highlighted notable sex differences in both osteoporosis diagnosis and physical activity. Women accounted for the vast majority (84.2%) of individuals diagnosed with osteoporosis, while only 15.8% were men. This difference was highly statistically significant and reinforces the well-established understanding that women are disproportionately affected by osteoporosis. Within the osteoporosis subgroup, sex was also linked to differences in activity levels. Specifically, 9.9% of men with osteoporosis met the high activity threshold, compared to only 4.6% of women. This difference was statistically significant, suggesting that men with osteoporosis are more likely to maintain higher levels of physical activity than women with the condition.

Finally, a logistic regression was conducted within the osteoporosis subgroup to identify predictors of high activity levels among those already diagnosed with the condition. This model was statistically significant and showed excellent fit, but only one variable, age group, was a significant predictor. As in the general population, individuals aged 70 and older were significantly less likely to engage in high levels of physical activity compared to those aged 40 to 55. Other factors, including sex, denture use, food insecurity, and cultural background, were not statistically significant within this group. Therefore, age stood out as the only valid predictor of high physical activity among individuals with osteoporosis.

In terms of comparing our findings with other research papers, we found out that our result about osteoporosis prevalence being higher among women is consistent with prior research. In a certain study, it states that osteoporosis disproportionately affects postmenopausal women due to estrogen decline (Cauley, 2003). Similarly, the finding that adults aged ≥70 had lower odds of meeting recreational activity guidelines mirrors previous work, which shows that physical activity typically declines with age due to comorbidities, reduced mobility, and social factors (Paterson & Warburton, 2010).

However, some findings diverged from expectations. For example, the lack of significant association between osteoporosis status and recreational activity in the full sample differs from studies that report significant relationships between osteoporosis status and physical activity. This may be due to our self-reported data, which can misclassify both activity and osteoporosis status. Additionally, threshold-based categorization (360 minutes/week) may mask more nuanced relationships (e.g., moderate vs. vigorous activity types) seen in other studies using accelerometry. This may have been a barrier and altered the results within our study. To explain the negative relationship between osteoporosis status and recreational activity, other papers have suggested that individuals with osteoporosis are often less physically active due to a combination of physical limitations, fear of injury, and lack of appropriate exercise guidance (Brooke-Wavell et al., 2022). According to Giangregorio et al. (2014), many individuals with low bone mass or osteoporosis limit their activity out of concern that physical movement may cause fractures, even though regular exercise is beneficial in reducing that risk.

In addition to psychological barriers, pain and mobility limitations following osteoporotic fractures, especially vertebral compression fractures, can significantly reduce one's ability or willingness to stay active. The National Institutes of Health (NIH, 2018) notes that people with osteoporosis may experience chronic pain and spinal deformities that make exercise more difficult, contributing to a decline in physical activity levels. Healthcare advice also plays a role. Many patients are unsure of which types of physical activity are safe, and in some cases, they may receive overly cautious recommendations that lead to complete avoidance of exercise (Osteoporosis Canada, 2020). This cautious behavior, while well-intentioned, can result in a sedentary lifestyle, further weakening bones and muscles over time.

As these results convey that osteoporosis status has a valid impact on the physical activity levels an individual would perform, we see a dissociation from our results which do not show a significant relationship.

Looking back at our study, we can see we had several key strengths that contributed to the quality and relevance of our findings. Firstly, we used data from the Canadian Community Health Survey (CCHS), a large, nationally representative, and population-based dataset. This allowed us to generalize our findings to Canadian adults aged 40 and older. We also believe our research addresses an important public health issue: osteoporosis, a highly prevalent chronic condition. By linking it to recreational physical activity, a modifiable behavior, we were able to highlight areas for potential intervention, especially in aging populations. We also conducted gender-stratified analyses to explore sex-based differences in activity levels, which added important nuance to the existing literature and helped address known disparities in osteoporosis management. In addition, we included multiple sociodemographic covariates such as age, marital status, sex, and ethnicity. This allowed us to interpret the relationship between osteoporosis and physical activity more holistically. From a methodological standpoint, we used appropriate statistical tests, including non-parametric methods for our non-normally distributed outcome variable and chi-square tests for categorical comparisons, ensuring that our inferences were valid. Most importantly, our findings, such as the lower rates of physical activity among women with osteoporosis, have practical, real-world implications for designing targeted interventions to support those most at risk.

That being said, our study also had several limitations. Since we used cross-sectional data, we were unable to determine causality. For example, we could not determine whether lower physical activity contributed to osteoporosis or whether this was a result of having the condition.

In addition, in terms of the way the data was collected, both osteoporosis status and physical activity levels were self-reported, which may have introduced recall bias or misclassification, particularly for participants who were undiagnosed or unaware of their condition. We were also limited in the detail of our activity data; recreational activity was measured only by total minutes per week, so we could not assess the type or intensity of the activity, which are crucial for osteoporosis management. Additionally, our sample was limited to residents of Alberta, which may affect how applicable our findings are to other regions of Canada with different demographics, climates, or healthcare access. We treated osteoporosis as a binary variable, without accounting for disease severity, fracture history, or treatment adherence, all of which could impact physical activity levels. Finally, some potentially influential factors like fear of falling, pain severity, and other comorbid health conditions (such as arthritis or depression) weren't included in the dataset, even though they likely played a role in shaping activity patterns in this population (Brooke-Wavell et al., 2022).

There are several valuable directions for future research that could build on the findings of our study. Firstly, longitudinal studies are essential to better understand how a diagnosis of osteoporosis influences physical activity patterns over time. While our research used cross-sectional data, which provides a snapshot at a single point, longitudinal data would allow us to track changes in behavior, attitudes, and health outcomes following diagnosis. This would help clarify whether reduced physical activity is a cause or a consequence of osteoporosis and could also help evaluate the long-term impact of interventions.

Secondly, qualitative research is needed to capture the lived experiences of individuals living with osteoporosis. While quantitative data (what we used) gives us important trends, it cannot fully uncover the emotional, psychological, and social dimensions of living with a chronic bone condition. In-depth interviews or focus groups could help uncover how individuals interpret their physical limitations, manage fear, and make decisions about activity. This is especially important in understanding gendered experiences, for instance, why women with osteoporosis may be less active than men, and how cultural or personal narratives around fragility and aging contribute to those differences.

Another important area for future investigation involves the psychological and physical barriers to physical activity, particularly pain and fear of falling. These are consistently cited in the literature as major deterrents to physical activity among people with osteoporosis, yet they are often underrepresented in quantitative survey datasets. Research that directly measures or addresses these barriers could help in designing more effective, supportive interventions.

Finally, we see strong value in evaluating the effectiveness of targeted, community-based interventions. Programs that are tailored based on gender, age, fitness level, or cultural background could help increase physical activity participation in individuals with osteoporosis. Testing these interventions through randomized controlled trials or community pilot programs could provide evidence-based strategies for improving health outcomes and promoting long-term adherence to physical activity.

Our findings also have several meaningful implications for healthcare practice and public health. For one, they can help healthcare providers better anticipate and address behavior changes that may follow an osteoporosis diagnosis. By understanding that some patients, especially women, may reduce their activity due to fear or uncertainty, providers can offer proactive counseling to encourage safe, regular movement rather than unintentionally reinforcing avoidance. This can be done throughout implementing gender-sensitive program development. If women with osteoporosis are participating less in physical activity, this calls for tailored interventions that address their unique concerns, such as offering female-only exercise groups, providing educational resources that build confidence, or engaging trusted healthcare professionals to dispel myths about fragility and movement.

At a broader level, these insights can inform public health policies aimed at reducing sedentary behavior and promoting active aging. Supporting physical activity among individuals with osteoporosis has the potential to improve not only bone health, but also overall well-being, functional independence, and quality of life. As the population ages and osteoporosis prevalence rises, these considerations will become increasingly important for sustainable, inclusive health systems.

References

Aoyagi, Y., & Shephard, R. J. (2013). Sex differences in relationships between habitual physical activity and health in the elderly: Practical implications for epidemiologists based on pedometer/accelerometer data from the Nakanojo Study. Archives of Gerontology and Geriatrics, 56(2), 327–338. https://doi.org/10.1016/j.archger.2012.11.006

Babiuch, A. S., Oestervemb, K., Lipińska, A., Stańczak, M. L., Cholewa, M., Makulec, K., Nowakowska, K., & Derengowska, M. H. (2021). Differences in the level of physical fitness and mobility among older women with osteoporosis and healthy women—Cross-sectional study. *Scientific Reports, 11*(1), Article 93483. https://doi.org/10.1038/s41598-021-93483-3

Brooke-Wavell, K., Skelton, D. A., Barker, K. L., Clark, E. M., De Biase, S., Arnold, S., Paskins, Z., Robinson, K. R., Lewis, R. M., Tobias, J. H., Ward, K. A., Whitney, J., & Leyland, S. (2022, May 16). *Strong, steady and straight: UK consensus statement on physical activity and exercise for osteoporosis*. British journal of sports medicine. https://pmc.ncbi.nlm.nih.gov/articles/PMC9304091/

Cauley, J. A. (2003). Effects of estrogen plus progestin on risk of fracture and bone mineral density: The Women's Health Initiative randomized trial. *JAMA, 290*(13), 1729–1738. https://doi.org/10.1001/jama.290.13.1729

Chen, Y., Luo, Z., Sun, Y., Zhou, Y., Han, Z., Yang, X., Kang, X., Lin, J., Qi, B., Lin, W., Guo, H., Guo, C., Go, K., Sun, C., Li, X., Chen, J., & Chen, S. (2022). The effect of denture-wearing on physical activity is associated with cognitive impairment in the elderly: A cross-sectional study based on the CHARLS database. *Frontiers in Neuroscience, 16*, Article 925398. https://doi.org/10.3389/fnins.2022.925398

Florvåg, A. G., Berg, Ø. A., Røksund, O. D., Jorem, G. T., & Bogen, B. E. (2025). Exercise interventions to improve bone mineral density in athletes participating in low-impact sports: A scoping review. *BMC Musculoskeletal Disorders, 26*(1), Article 8316. https://doi.org/10.1186/s12891-025-08316-5

Gammage, K. L., Gasparotto, J., Mack, D. E., & Klentrou, P. (2011). Gender differences in osteoporosis health beliefs and knowledge and their relation to vigorous physical activity in university students. Journal of American College Health, 60(1), 58–64. https://doi.org/10.1080/07448481.2011.570399

Howe, T. E., Shea, B., Dawson, L. J., Downie, F., Murray, A., Ross, C., Harbour, R. T., Caldwell, L. M., & Creed, G. (2011). Exercise for preventing and treating osteoporosis in postmenopausal women. Cochrane Database of Systematic Reviews, 2011(7), Article CD000333. https://doi.org/10.1002/14651858.cd000333.pub2

Kelley, C. (2018, January 1). *Modes of physical activity reported by older adults living with osteoporosis in Canada*. Academia.edu. https://www.academia.edu/92067107/Modes_of_physical_activity_reported_by_older_adults_living_with_osteoporosis_in_Canada

McArthur, C., Lee, A., Alrob, H. A., Adachi, J. D., Giangregorio, L., Griffith, L. E., Morin, S., Thabane, L., Ioannidis, G., Lee, J., Leslie, W. D., & Papaioannou, A. (2022). An update of the prevalence of osteoporosis, fracture risk factors, and medication use among community-dwelling older adults: Results from the Canadian Longitudinal Study on Aging (CLSA). Archives of Osteoporosis, 17(1), 1–11. https://doi.org/10.1007/s11657-022-01073-1

Meyer, F., König, H., & Hajek, A. (2019). Osteoporosis, fear of falling, and restrictions in daily living: Evidence from a nationally representative sample of community-dwelling older adults. Frontiers in Endocrinology, 10, Article 646. https://doi.org/10.3389/fendo.2019.00646

Morin, S. N., Feldman, S., Funnell, L., Giangregorio, L., Kim, S., McDonald-Blumer, H., Santesso, N., Ridout, R., Ward, W., Ashe, M. C., Bardai, Z., Bartley, J., Binkley, N., Burrell, S., Butt, D., Cadarette, S. M., Cheung, A. M., Chilibeck, P., Dunn, S., ... Wark, J. D. (2023). Clinical practice guideline for management of osteoporosis and fracture prevention in Canada: 2023 update. Canadian Medical Association Journal, 195(39), E1333–E1348. https://doi.org/10.1503/cmaj.221647

Nguyen, V. H. (2016). Osteoporosis prevention and osteoporosis exercise in community-based public health programs. Osteoporosis and Sarcopenia, 3(1), 18–31. https://doi.org/10.1016/j.afos.2016.11.004

Pinheiro, M. B., Oliveira, J., Bauman, A., Fairhall, N., Kwok, W., & Sherrington, C. (2020). Evidence on physical activity and osteoporosis prevention for people aged 65+ years: A systematic review to inform the WHO guidelines on physical activity and sedentary behaviour. International Journal of Behavioral Nutrition and Physical Activity, 17(1), 1–15. https://doi.org/10.1186/s12966-020-01040-4

Paterson, D. H., & Warburton, D. E. (2010). Physical activity and functional limitations in older adults: a systematic review related to Canada's Physical Activity Guidelines. International Journal of Behavioral Nutrition and Physical Activity, 7(1), 38. https://doi.org/10.1186/1479-5868-7-38

Simmonds, B. A. J., Hannam, K. J., Fox, K. R., & Tobias, J. H. (2015). An exploration of barriers and facilitators to older adults' participation in higher impact physical activity and bone health: A qualitative study. Osteoporosis International, 27(3), 979–987. https://doi.org/10.1007/s00198-015-3376-7

Stanghelle, B., Bentzen, H., Giangregorio, L., Pripp, A. H., Skelton, D. A., & Bergland, A. (2020). Physical fitness in older women with osteoporosis and vertebral fracture after a resistance and balance exercise programme: 3-month post-intervention follow-up of a randomised controlled trial. BMC Musculoskeletal Disorders, 21(1), 1–10. https://doi.org/10.1186/s12891-020-03495-9

Su, Y., Chen, Z., & Xie, W. (2020). Swimming as treatment for osteoporosis: A systematic review and meta-analysis. BioMed Research International, 2020, Article 6210201. https://doi.org/10.1155/2020/6210201

Vásquez, E., Shaw, B. A., Gensburg, L., Okorodudu, D., & Corsino, L. (2013). Racial and ethnic differences in physical activity and bone density: National Health and Nutrition Examination Survey, 2007–2008. Preventing Chronic Disease, 10. https://doi.org/10.5888/pcd10.130183

Appendix

Supplementary Table 2. Logistic Regression Predicting High Recreational Activity Levels (≥360 min/week) in the Full Sample (N = 6870)

Results of a logistic regression model examining the relationship between demographic and health variables and the likelihood of engaging in at least 360 minutes of recreational activity per week. Categorical variables (e.g., age group, cultural background, marital status) were entered as factors. The reference groups are: age 40-55, white, not wearing dentures, and food security. Odds ratios (Exp(B)) with 95% confidence intervals are reported.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	Age			17.286	2	<.001			
	Age(1)	-.194	.105	3.390	1	.066	.824	.670	1.013
	Age(2)	-.676	.163	17.240	1	<.001	.509	.370	.700
	Has osteoporosis(1)	.085	.186	.208	1	.649	1.088	.756	1.567
	Wears dentures, dental prosthesis, false teeth(1)	.393	.131	8.981	1	.003	1.482	1.146	1.916
	Cultural/Racial background(1)	-.582	.184	9.970	1	.002	.559	.389	.802
	Food Insecurity			4.394	2	.111			
	Food Insecurity(1)	-.717	.482	2.217	1	.136	.488	.190	1.255
	Food Insecurity(2)	-.142	.401	.126	1	.723	.867	.395	1.904
	Marital Status			1.860	3	.602			
	Marital Status(1)	-.228	.210	1.179	1	.277	.796	.528	1.201
	Marital Status(2)	-.080	.119	.451	1	.502	.923	.732	1.165
	Marital Status(3)	-.132	.154	.727	1	.394	.877	.648	1.186
	Constant	-2.489	.453	30.232	1	<.001	.083		

a. Variable(s) entered on step 1: Age, Has osteoporosis, Wears dentures, dental prosthesis, false teeth, Cultural/Racial background, Food Insecurity, Marital Status.

Supplementary Table 3. Logistic Regression Predicting High Recreational Activity Levels (≥360 min/week) Among Participants with Osteoporosis (N = 640)

Logistic regression results examining predictors of high recreational activity (≥360 minutes/week) among individuals diagnosed with osteoporosis. Age was the only significant predictor in this model, with participants aged 70 and older having lower odds of meeting activity guidelines compared to those aged 40-55 (p = .036). Categorical predictors (e.g., age group, sex, cultural background) were entered as factors. Reference categories: age 40-55, White, not wearing dentures, food security. Odds ratios (Exp(B)) with 95% confidence intervals are reported.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	Age			4.395	2	.111			
	Age(1)	-.713	.525	1.850	1	.174	.490	.175	1.370
	Age(2)	-1.373	.655	4.390	1	.036	.253	.070	.915
	Wears dentures, dental prosthesis, false teeth(1)	-.419	.425	.974	1	.324	.658	.286	1.512
	Cultural/Racial background(1)	-.726	1.077	.455	1	.500	.484	.059	3.992
	Food Insecurity			.000	2	1.000			
	Food Insecurity(1)	-.696	11574.651	.000	1	1.000	.498	.000	.
	Food Insecurity(2)	17.658	9892.573	.000	1	.999	46655841.2	.000	.
	Marital Status			2.111	3	.550			
	Marital Status(1)	-1.101	1.110	.984	1	.321	.333	.038	2.928
	Marital Status(2)	-.558	.478	1.364	1	.243	.572	.224	1.460
	Marital Status(3)	-.236	.640	.136	1	.712	.790	.225	2.771
	Sex	-.587	.433	1.833	1	.176	.556	.238	1.300
	Alternate physical activity indicator	-17.616	1463.538	.000	1	.990	.000	.000	.
	Constant	.464	10000.247	.000	1	1.000	1.590		

a. Variable(s) entered on step 1: Age, Wears dentures, dental prosthesis, false teeth, Cultural/Racial background, Food Insecurity, Marital Status, Sex, Alternate physical activity indicator.

Unhealthy Coping Mechanisms: Assessing Rates of Alcohol Consumption in Canadian Cancer Patients Aged 30 to 60

Shruthi Gundala, Sakti Putkaran, Elizabeth Tsyben

McMaster University

Department of Interdisciplinary Science

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8, 2025

Unhealthy Coping Mechanisms: Assessing Rates of Alcohol Consumption in Canadian Cancer Patients Aged 30 to 60

Introduction

Surviving with the burden and stress of a chronic disease such as cancer often leads patients to become reliant on unhealthy substance-based coping behaviours (Wu et al., 2018). While physicians care for the physical health of chronic disease patients via targeted treatment plans, it is also important to care for the patient's mental health. The World Health Organization defines "health" as not only encompassing one's physical state, but also one's mental well-being (World Health Organization, 2024). As such, the maintenance of good health also implies preserving one's *mental health*. One study found that the patient's perception of the illness and the coping mechanisms they develop correlates with their mental health, which can consequently affect their physical health (Dempster et al., 2015).

Yet, the stigma against mental health disorders is still pervasive among medical professionals, thus negatively affecting the recovery process of cancer patients (Crapanzano et al., 2023). Hack and associates found that physicians spend an insufficient amount of time addressing cancer patients' psychological concerns (2009). Furthermore, access to adequate professional psychological support even for cancer patients may be limited by a lack of appointment vacancies and specific empirically-supported therapies (Walshe et al., 2017).

The minimal importance given by physicians to address the chosen coping strategies of cancer patients enables patients to develop unhealthy thought patterns (Drageset et al., 2016; Levy & Cartwright, 2015). This opens the door for psychological comorbidities in chronic illness patients, most commonly depression, anxiety or both as a result of adjustment disorder (Chaturvedi, 2012). The presence of psychological disorders is quite common in cancer patients, with a study from 2013 reporting around a self-reported 12 to 25% prevalence of anxiety and 5 to >40% prevalence of depression, with many patients not receiving adequate care for these comorbidities (Grassi et al., 2023). However, oftentimes physicians do not use specific diagnostic language pertaining to mental health disorders (Yusufov et al., 2019). Additionally, with many studies conducted in the United States, there is minimal empirical literature surrounding Canadian cancer patients' development of negative coping mechanisms.

Given the prevalence of psychological disorders in cancer patients, substance usage (specifically alcohol) is a common method of pain management (Yusufov et al., 2019). Alcohol intake is linked with liver, breast and stomach cancer to name a few, and accounts for 4% of all cancer deaths in the United States (American Cancer Society, 2020). Turning to global statistics, one European study reported that hazardous drinking was prevalent in 20% of cancer survivors 50 years or older (Shi et al., 2023). Statistics Canada corroborates these findings, reporting 18-34 year olds as the heaviest drinkers (males and females) and 35-49 year olds as the second highest cohort of heavy drinkers (2019). The United States Centers for Disease Control and Prevention defines heavy drinking as 8 or more drinks consumed in a week for women, and 15 drinks a week for men (Centers for Disease Control, 2024). One study in the United States found that 38.3% of cancer patients engaged in what the researchers defined as hazardous drinking habits (Shi et al., 2023). The *Diagnostic and Statistical Manual of Mental Disorders* (DSM) generally defines consuming more alcohol than planned, craving the substance, experiencing withdrawal, building a tolerance to alcohol and repeated use of the substance in dangerous situations as symptoms of Substance Use Disorders (SUDs) (American Psychiatric Association, 2013). Bearing this in mind, Yusufov and associates report that patients with cancer report a much higher rate of SUDs at 35% in contrast with the general population (8.4%) (2019).

Not only is alcohol a risk factor for cancers, its consumption to cope with the mental stress and fatigue of cancer is not an effective nor healthy remedy. This study will address the gap of a lack of Canadian literature on this topic, and the focus on 30-60 year old cancer patients includes a known cohort of heavy drinkers. Moreover, the exploration of sex differences in perceived mental health of cancer patients is not commonly explored in empirical literature, specifically when pertaining to investigating possible risk factors for substance dependence. Thus, the focus of this study is on active cancer patients (rather than previous patients) and the damaging coping mechanism of consuming a large amount of alcohol. The goal is to investigate if the prevalence of self-reported high alcohol consumption determined as over 8 drinks a week for women and over 15 drinks a week for men in the past year is different across 30 to 60-year-old Canadians with an active clinically-determined cancer diagnosis compared to healthy individuals in the same age group. Additionally, the association between the sex of 30 to 60-year-old Canadian patients with an active clinically-determined cancer diagnosis and their perceived (as determined by self-reported data) as well as actual mental health (as determined by medication taken by the patient and care type received) in the past year (2013) will also be examined.

Methods

This observational, cross-sectional study was conducted using data from the annual Canadian Community Health Survey (CCHS) from 2011-2012, accessed through the data repository Odesi. The CCHS was our chosen dataset as it collects information related to health status and health care utilization, receiving responses from about 98% of the Canadian population aged 12 and over, thus providing our study with a large and diverse sample. It provides data on Canadians living in all provinces and territories via survey responses to questions regarding both their physical and mental health.

Among the 124,929 respondents to the 2011-2012 survey, inclusion and exclusion criteria were set in order to minimize the number of study participants within the database data from which would be used. Thus, individuals aged 30-60 who were either healthy or diagnosed with cancer were included as study participants, in our control and exposure group, respectively. Participants were categorized into three groups for ease of data analysis: 30 to 39, 40 to 49 and 50 to 59 years of age.

Individuals with a past or currently inactive cancer diagnosis were excluded, along with those who did not meet the threshold set for excessive drinking, which was defined as 8 drinks/week for females and 15 drinks/week for males. Respondents with necessary data missing from any of these categories were also removed from the dataset, thus resulting in a final study population of 3,219 cases.

The primary outcome of this study was to investigate possible differences in alcohol consumption between cancer patients and healthy individuals. To do so, data from the survey asking participants about their “self-reported weekly alcohol consumption” was analyzed, where participants responded by providing numerical values, increasing the validity of our data. This methodology was adapted from Shi and associates, who used categorical data from participants regarding their alcohol consumption in the past year on a monthly or less, or weekly basis (2023).

The secondary outcome was to examine the perceived mental health of cancer patients and healthy individuals, since this is often related to one’s alcohol consumption or substance usage (Yusufov et al., 2019). To assess this, data was compiled from a CCHS survey question asking participants to rate their "perception of [their] mental health", on a scale from 1 to 5, where 1 = excellent, 2 = very good, 3 = good, 4 = fair and 5 = poor. This metric was chosen in order to highlight any negative emotions in individuals who consume excessive alcohol, similar to the use of interviews to assess a patient's mindset in relation to coping strategies (Walshe et al., 2017). Other variables considered when addressing these outcomes include self-reported weight, to ensure the general health of study participants, as well as demographic information such as sex and geography (i.e. Canadian province or territory).

All data cleanup and statistical analyses were conducted using IBM SPSS Statistics Version 30.0.0, with significance set to $\alpha = 0.05$. As per Table 1, all data is categorized by healthy participants and individuals with a current clinical cancer diagnosis, with demographic data (consolidated as categorical) summarized by individuals' response to the CCHS survey. To test for normality in the collected continuous data (i.e. the number of drinks consumed by participants weekly), a Shapiro-Wilk test was used. From this, values of kurtosis and skewness were analyzed to further determine the normality of the dataset. Following this, a Mann-Whitney U test was conducted to evaluate for significant differences between control and exposure groups. Additionally, a biserial correlation analysis was conducted, with Spearman's Rho evaluated to further test for possible significant relationships.

Results

The dataset was first manipulated to only include cases within the target age range of 30-60 years old, removing individuals between 12 - 29 and 60 - 65+. Missing cases were then identified for alcohol consumption, cancer diagnosis status, and all necessary demographic descriptors. This reduced the original dataset from 124,929 cases to 25,976. The dataset was reduced to 3,219 after deleting all cases of women drinking under 8 drinks and men drinking under 15 drinks. After narrowing the dataset down, age, sex, geographic, BMI, and mental health differences were assessed across healthy individuals and cancer patients, as reported in Table 1.

Table 1.

Characteristics of census population across cancer patients and healthy participants.

	Cancer (n = 43)	**Healthy** (n = 3176)	**Total** (n = 3219)
Age			
30 to 39 years	3 (6.98)	920 (28.97)	923 (28.67)
40 to 49 years	6 (13.95)	947 (28.82)	953 (29.61)
50 to 59 years	34 (79.07)	1309 (42.21)	1343 (41.72)
Sex			
Female	28 (65.11)	1628 (51.26)	1656 (51.44)
Male	15 (34.89)	1548 (48.74)	1563 (48.56)
Geography			
Newfoundland & Labrador	2 (4.65)	151 (4.75)	153 (4.75)
Quebec	18 (41.86)	1022 (32.18)	1040 (32.31)
Ontario	20 (46.51)	1556 (48.99)	1576 (48.96)
Manitoba	0	237 (7.46)	237 (7.36)
Saskatchewan	3 (6.98)	210 (6.62)	213 (6.62)
Self-Reported Weight			
Underweight	0	43 (1.35)	43 (1.34)
Normal Weight	17 (39.53)	1396 (43.95)	1413 (43.89)
Overweight	16 (37.21)	1165 (36.69)	1181 (36.69)
Obese	10 (23.26)	572 (18.01)	582 (18.08)
Self-Reported Mental Health			
Excellent	13 (30.23)	1052 (33.12)	1065 (33.08)
Very Good	17 (39.53)	1205 (37.94)	1222 (37.96)
Good	8 (18.61)	700 (22.04)	708 (21.99)
Fair	3 (6.98)	174 (5.48)	177 (5.50)

Poor	2 (4.65)	45 (1.42)	47 (1.47)

Note. Values are reported as n(%). Self-reported weight was determined by the variable's self-determined BMI. To be qualified as a cancer patient, participants required a current cancer diagnosis as determined by claiming 'NOT APPLICABLE' to the 'Have you ever been diagnosed with cancer?' portion of the census. Healthy patients include those with no cancer diagnosis and those with past cancer diagnosis

Statistical analysis

The only continuous variable used for analysis was self-reported weekly alcohol consumption. To ensure normality, a Shapiro-Wilk test was run. The resultant p-value was <0.001, indicating a significant difference from normal distribution. This was further confirmed by positive values for kurtosis (31.686) and skewness (4.245).

A Mann-Whitney U test was conducted to measure whether between-group differences were significant. The asymptotic significance value reported was 0.278, confirming the retention of the null hypothesis. This indication confirmed that there was no significant difference in the medians across alcohol consumption between individuals with and without cancer. To further assess group differences, a biserial correlation analysis was run. Spearman's Rho for this relationship was $\rho = 0.019$, indicating a potential very weak correlation. The resultant p-value, however, was 0.278, indicating no correlation between cancer status and alcohol consumption. A visual representation of this relationship can be seen in Figure 1. Men with cancer had the highest upper quartile range for alcohol consumption, while no difference is seen across groups in females. Healthy groups across both genders had a higher maximum value, likely imposed by the quantity of outliers and extreme values.

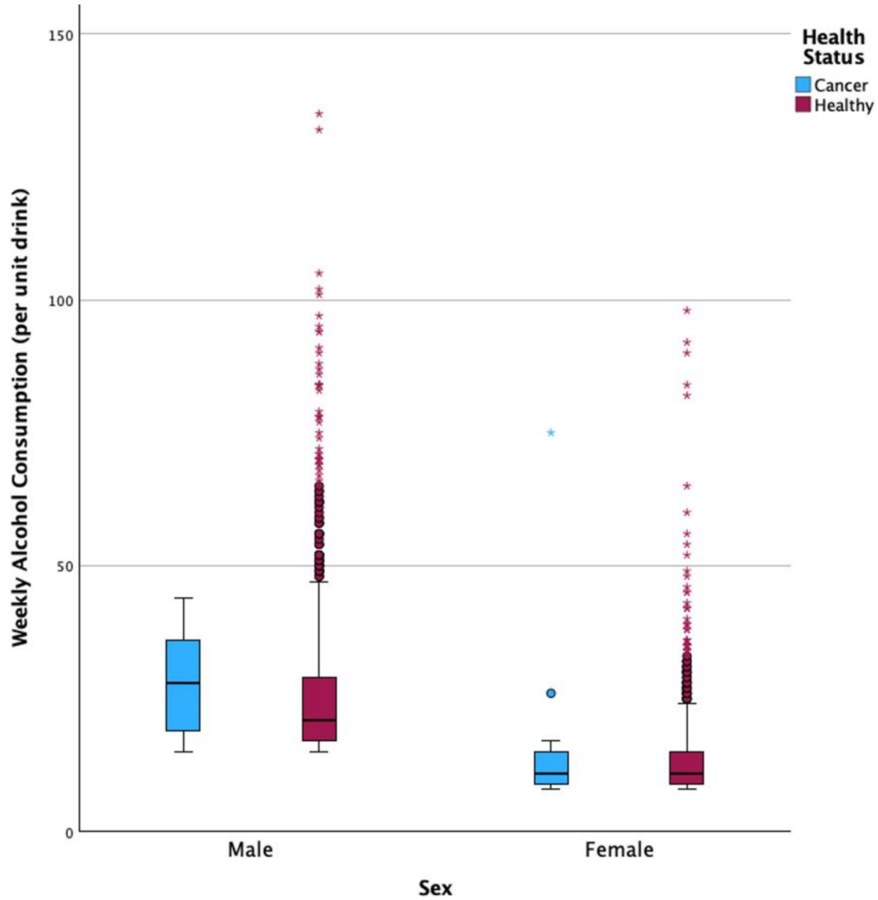


Figure 1. Self-reported weekly alcohol consumption across healthy individuals and those with a cancer diagnosis. Alcohol consumption - as reported by the number of unit drinks per week - was compared across health status and sex. Only cases of females reporting 8 or more drinks per week (n = 1656) and males reporting 15 or more (n = 1563) were retained for analysis. Statistical analysis was performed by obtaining Spearman's rank correlation coefficient, $\rho = 0.019$, $p = 0.278$.

A secondary analysis of the effect of a cancer diagnosis on self-perceived mental health was conducted. A bivariate correlation analysis was run, producing a Rho value of -0.009. This very weak negative correlation was insignificant ($p = 0.604$), thereby confirming no relationship between cancer presence and self-perceived mental health. Table 2 depicts a brief listing summary of all statistical analysis results obtained.

Table 2.
Summary of statistical analyses and corresponding outputs.

Primary analysis: Relationship between alcohol consumption and cancer presence	
Shapiro-Wilk	<0.001*
Mann-Whitney U	0.278
Spearman's Rho	0.019
p	0.278
Secondary Analysis: Relationship between cancer presence and self-perceived mental health	
Spearman's Rho	-0.009
p	0.604

Discussion

This study aimed to investigate the prevalence of high alcohol consumption among 30 to 60 year old Canadians with an active clinically determined cancer diagnosis, compared to healthy individuals. Our findings showed that there is no significant difference in alcohol consumption between individuals with an active clinically determined cancer diagnosis, and healthy individuals. Similarly, our secondary research outcome investigated cancer patients' perceived mental health and alcohol consumption, where results through a bivariate correlation showed that there is once again no relationship between the presence of cancer and self-perceived mental health in individuals.

Prior to the results being conducted we hypothesized that individuals with an active cancer diagnosis would consume higher rates of alcohol due to various psychological factors or as a coping mechanism. However, the Shapiro-Wilk test, Mann Whitney U test and the Spearman’s Rho test did not support the hypothesis. They indicated a weak and non-significant relationship between cancer diagnosis and alcohol consumption. These negative and non-significant results show the different and complex patterns in populations of cancer patients and indicate that cancer alone may not be a strong enough predictor of alcohol use among individuals. Additionally, there was also no significant relationship found between self-perceived mental health and alcohol use between individuals. These may be due to several factors including the different cancer types and stages of different individuals, differences in support systems that individuals may be receiving, and differences in coping mechanisms of individuals. Although these results did not align with our expected results, they still offer insight on considering different psychological variables and coping mechanisms among cancer patients.

Moreover, these findings contrast with other pre-existing literature. A 2023 study in the All of Us research program found that 77.7% of cancer survivors in the US reported alcohol consumption, with over one third of them engaging in hazardous drinking while undergoing active cancer treatment (Shi et al., 2023). Similarly, findings from the National Cancer Institute showed that many cancer survivors continue to consume alcohol despite the known risks linked to new or the recurrence of cancer (Phillips, 2023).

There are several reasons for these differences including methodological differences as well as differences in healthcare systems, and cultural norms. Healthcare differences between the US and Canada may influence the mental health and support services that cancer patients receive, ultimately influencing the reliance of substances like alcohol as coping mechanisms. Additionally, the definitions of the term hazardous and excessive may have differences in a survey design, influencing the findings. In the All of Us research study, a more broader age range was included, as well as self-reported pre-existing cancer diagnoses. Both these studies heavily relied on self-reported data, which may be subject to recall bias. The study also did not specify the cancer type or stage, which may have confused early versus late stage experiences of cancer, and impact alcohol consumption.

Our study provides insights specific to the Canadian population, since most existing literature focuses on American and European data. By focusing on individuals with an active cancer diagnosis, rather than cancer survivors, our study highlights the use of alcohol as a coping mechanism throughout an individual's cancer journey. Moreover, our findings emphasize the need for a multifaceted approach at analysing alcohol consumption, by not just considering medical factors but also different psychological, social and economic factors that may have an influence. Moving forward, this research can help make informed healthcare policies and interventions aimed at improving the overall health and well-being of cancer patients in Canada.

Our study also had several limitations that should be considered. Firstly, due to the cross-sectional design of the study it is hard to establish causality, or different alcohol patterns that may be present before a cancer diagnosis. It is unclear whether drinking patterns seen are due to a cancer diagnosis, or were pre-existing patterns that were established before. In order to assess this accurately, a longitudinal study would be needed to accurately assess how alcohol evolved before, during and after treatment. Secondly, the self-reported and self-perceived data introduces potential bias. Participants may under report data, not remember, or intentionally downplay their drinking behaviour, especially if they are worried about others' perceptions. The dataset also did not account for different types of cancer, stages of cancer, or the time since cancer diagnosis, which could all influence drinking behaviours and coping mechanisms in individuals. Furthermore, we did not account for any confounding variables in our study such as socioeconomic status, psychological distress or social support systems which could all impact alcohol consumption. Although our statistical analysis provided insightful results and relied on non-parametric tests, the weak correlation suggests that future studies could consider additional factors and analyses to account for potential confounders. Finally, our study lacked a true control group. While we did compare active clinically determined cancer patients to healthy patients, we had no control over other unmeasured health conditions or life factors that may influence alcohol use in individuals.

Overall, while our findings showed no strong relationship between cancer diagnosis and alcohol consumption, or between self-perceived mental health in cancer patients and healthy patients, it emphasizes the need to explore other complex social, psychological or even clinical factors that may influence substance use in cancer patients, as studies have shown that individuals with general life stressors are more prone to alcohol use (Keyes et al., 2015). Rather than relying on self-perceived mental health measures, using more comprehensive assessments and validated screening tools may provide more accurate results without too much bias. Further, looking at patients on a case by case basis may provide details on personal aspects that may affect alcohol consumption in individuals that are hard to identify through a survey. Public health initiatives can also emphasize routine mental health assessments for cancer patients to identify and mitigate substance use risks early on. Encouraging healthcare providers to routinely screen for alcohol use in patients and implementing targeted interventions can help patients develop healthier coping mechanisms (McKnight-Eily et al., 2017). This will also help address underreporting in patients and use physician data for more accurate results. Additionally, improving access to psychological and mental health screenings for cancer patients may reduce the reliance of maladaptive coping strategies and contribute to a more holistic approach to patient care.

References

American Cancer Society. (2020, June 9). *Alcohol Use and Cancer*. Www.cancer.org; American Cancer Society. https://www.cancer.org/cancer/risk-prevention/diet-physical-activity/alcohol-use-and-cancer.html

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

CDC. (2024, May 9). *Alcohol Use and Your Health*. Alcohol Use. https://www.cdc.gov/alcohol/about-alcohol-use/index.html

Chaturvedi, S. K. (2012). Psychiatric oncology: Cancer in mind. *Indian Journal of Psychiatry*, *54*(2), 111–118. https://doi.org/10.4103/0019-5545.99529

Crapanzano, K. A., Deweese, S., Pham, D., Le, T., & Hammarlund, R. (2023). The role of bias in clinical decision-making of people with serious mental illness and medical co-morbidities: A scoping review. *The Journal of Behavioral Health Services & Research*, *50*(2). https://doi.org/10.1007/s11414-022-09829-w

Dempster, M., Howell, D., & McCorry, N. K. (2015). Illness perceptions and coping in physical health conditions: A meta-analysis. *Journal of Psychosomatic Research*, *79*(6), 506–513. https://doi.org/10.1016/j.jpsychores.2015.10.006

Drageset, S., Lindstrøm, T. C., & Underlid, K. (2016). “I just have to move on”: Women’s coping experiences and reflections following their first year after primary breast cancer surgery. *European Journal of Oncology Nursing*, *21*, 205–211. https://doi.org/10.1016/j.ejon.2015.10.005

Grassi, L., Caruso, R., Riba, M. B., Lloyd-Williams, M., Kissane, D., Rodin, G., McFarland, D., Campos-Ródenas, R., Zachariae, R., Santini, D., & Ripamonti, C. I. (2023). Anxiety and depression in adult cancer patients: ESMO Clinical Practice Guideline. *ESMO Open*, *8*(2), 101155. https://doi.org/10.1016/j.esmoop.2023.101155

Hack, T. F., Pickles, T., Ruether, J. D., Weir, L., Bultz, B. D., & Degner, L. F. (2009). Behind closed doors: systematic analysis of breast cancer consultation communication and predictors of satisfaction with communication. *Psycho-Oncology*, *19*(6), 626–636. https://doi.org/10.1002/pon.1592

Keyes, K., Hatzenbuehler, M., Grant, B. F., & Hasin, D. S. (2015). Stress and Alcohol: Epidemiologic Evidence. *Alcohol Research : Current Reviews*, *34*(4), 391. https://pmc.ncbi.nlm.nih.gov/articles/PMC3797525/

Levy, A., & Cartwright, T. (2015). Men's strategies for preserving emotional well-being in advanced prostate cancer: An interpretative phenomenological analysis. *Psychology & Health*, *30*(10), 1164–1182. https://doi.org/10.1080/08870446.2015.1040016

McKnight-Eily, L. R., Henley, S. J., Green, P. P., Odom, E. C., & Hungerford, D. W. (2017). Alcohol Screening and Brief Intervention: A Potential Role in Cancer Prevention for Young Adults. *American Journal of Preventive Medicine*, *53*(3), S55–S62. https://doi.org/10.1016/j.amepre.2017.04.021

Phillips, C. (2023, September 15). *In People with Cancer, Heavy Drinking is Common - NCI*. Www.cancer.gov. https://www.cancer.gov/news-events/cancer-currents-blog/2023/cancer-survivors-alcohol-drinking-common

Scoccianti, C., Lauby-Secretan, B., Bello, P.-Y., Chajes, V., & Romieu, I. (2014). Female Breast

Cancer and Alcohol Consumption: A Review of the Literature. *American Journal of Preventive Medicine*, *46*(3, Supplement 1), S16–S25. https://doi.org/10.1016/j.amepre.2013.10.031

Shi, M., Luo, C., Oduyale, O. K., Zong, X., LoConte, N. K., & Cao, Y. (2023). Alcohol Consumption Among Adults With a Cancer Diagnosis in the All of Us Research Program. *JAMA Network Open*, *6*(8), e2328328. https://doi.org/10.1001/jamanetworkopen.2023.28328

Statistics Canada. (2019, June 25). *Heavy Drinking, 2018*. Statcan.gc.ca; Government of Canada, Statistics Canada. https://www150.statcan.gc.ca/n1/pub/82 625-x/2019001/article/00007-eng.htm

Walshe, C., Roberts, D., Appleton, L., Calman, L., Large, P., Lloyd-Williams, M., & Grande, G. (2017). Coping Well with Advanced Cancer: A Serial Qualitative Interview Study with Patients and Family Carers. *PLOS ONE*, *12*(1), e0169071. https://doi.org/10.1371/journal.pone.0169071

World Health Organization. (2024). *Constitution of the World Health Organization*. World Health Organization; World Health Organization. https://www.who.int/about/governance/constitution

Wu, L.-T., Zhu, H., & Ghitza, U. E. (2018). Multicomorbidity of chronic diseases and substance use disorders and their association with hospitalization: Results from electronic health records data. *Drug and Alcohol Dependence*, *192*, 316–323. https://doi.org/10.1016/j.drugalcdep.2018.08.013

Yusufov, M., Braun, I. M., & Pirl, W. F. (2019). A systematic review of substance use and substance use disorders in patients with cancer. *General Hospital Psychiatry*, *60*, 128–136. https://doi.org/10.1016/j.genhosppsych.2019.04.016

Investigating the Impact of Sleep Disturbances on Emotional Well-being and Positive Functioning Among Canadian Adults: A Cross-Sectional Study

Harsha Arunagiri, Zoë Brandt, Abisha Vasudevan

McMaster University

Department of Interdisciplinary Science

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8th, 2025

Investigating the Impact of Sleep Disturbances on Emotional Well-being and Positive Functioning Among Canadian Adults: A Cross-Sectional Study

Introduction

Recent literature examines the correlation between poor mental health and psychological well-being on unhealthy aging, poor life satisfaction, and various physical, mental, and social disadvantages (Cross et al., 2018; Kubzansky et al., 2015; Kushlev et al., 2020; Ngamaba et al., 2017). Sleep is widely recognized as a fundamental pillar of positive mental health and wellness (Clement-Carbonell et al., 2021). Research has consistently shown that sleep is essential not only for memory consolidation but also for emotional processing (Paller et al., 2022). In fact, Walker and van der Helm (2009) note that disruptions in sleep can impair cognitive function and lead to heightened emotional distress, underscoring the role of sleep in maintaining optimal brain function. Given these critical roles, sleep disturbances have increasingly been recognized as a public health concern, particularly as they relate to the exacerbation of negative emotional states (Chattu et al., 2018).

Sleep disturbances are also a major public health concern in Canada (Clayborne et al., 2017). Recent reports indicate that nearly one-third of Canadian adults experience inadequate sleep, placing a significant burden on public health systems (Clayborne et al., 2023). This prevalence highlights not only the magnitude of the problem but also the potential for widespread impacts on both mental and physical health across the population (Kirkbride et al., 2024). Despite these insights, it seems that literature has yet to fully address how the frequency of sleep disturbances over an extended period may directly affect reported emotional well-being and positive functioning.

Recent studies have begun to unravel the complex, bidirectional relationship between sleep and emotional well-being (Franzen & Buysse, 2008; Goldstein & Walker, 2014; Philbrook

& Macdonald-Gagnon, 2021). In a study of college students, Philbrook and Macdonald-Gagnon (2021) found that the interplay between sleep and emotional distress is influenced by social factors such as loneliness. Their findings suggest that poor sleep quality may not only result from emotional distress but also exacerbate it, creating a cyclical pattern of negative outcomes (Philbrook & Macdonald-Gagnon, 2021). Similarly, Franzen and Buysse (2008) investigated the relationship between sleep disturbances and depression, highlighting that sleep problems can be both a precursor to and a consequence of depressive symptoms. This research emphasizes that sleep disturbances are not merely an isolated symptom but are closely intertwined with broader emotional health concerns (Franzen & Buysse, 2008). In addition to these psychological dimensions, sleep disturbances have been linked to impairments in emotional regulation. Goldstein and Walker (2014) discuss how inadequate sleep can lead to heightened stress reactivity and mood instability, suggesting that individuals who experience chronic sleep disruptions may be at risk for poorer emotional regulation. These findings have significant implications, as emotional regulation is a key component of positive functioning and overall quality of life.

Positive functioning, which is defined by resilience, optimism, life satisfaction, and adaptive emotional regulation, is an outcome that is not only the absence of negative affect but also the presence of positive mental health (Gautam et al., 2024). Tout et al.'s 2023 study assesses positive psychological traits, such as emotional regulation, and their correlation with sleep quantity and quality (Tout et al., 2023). Their findings indicate that positive psychological traits may help improve sleep quantity and quality (Tout et al., 2023). In addition, a study by Sin et al., (2017) found that nights of better sleep were correlated with higher positive functioning and reduced odds of stressors (Sin et al., 2017). These findings suggest there is a bidirectional relationship between positive functioning and sleep quality which is similar to that of emotional well-being and sleep quality.

Closer examination of literature reveals that while multiple studies have explored the link between sleep quality and mental health, it appears that few have investigated the role of sleep disturbances in shaping emotional well-being and positive functioning (Franzen & Buysse, 2008; Goldstein & Walker, 2014). Furthermore, research has often been limited to younger (e.g. college students) or less ethnically diverse demographic groups (mainly Caucasian), leaving a void in our understanding of how these processes unfold within a more diverse adult population (Bibbins-Domingo & Helman, 2022). Given the significant burden of sleep disturbances in Canada (Clayborne et al., 2023)., there is a pressing need to explore these associations among a broader adult population. To address this gap, this study examines whether self-reported sleep disturbances over the past 10 years are associated with self-reported emotional well-being and positive functioning measured by self-reported Likert scale survey distributed among Canadian adults aged 20 and older, compared to their counterparts who do not experience sleep disturbances. This issue is important in cultivating a greater understanding of overall well-being in adults, alleviating negative emotional states and enhancing positive mental health indicators.

Methods

Study Design and Database

This study used a cross-sectional observational design to investigate correlations between self-reported sleep disturbances and mental health outcomes among Canadian adults over the age of 20. Data was collected from the 2022 Mental Health and Access to Care Survey (MHACS), a nationally representative database administered by Statistics Canada between March 17 to July 31, 2022, and accessed through Odesi (Statistics Canada, 2024). The MHACS used a stratified simple random sampling approach and was designed to assess mental health status, access to care services, and the impact of the COVID-19 pandemic over the past decade (Statistics Canada, 2024). Data collection occurred via computer-assisted telephone interviews presented in Canada's official languages, English and French. This method reduced respondent burden and maintained consistency with prior iterations of the Canadian Community Health Survey and MHACS that used an electronic questionnaire application.

Study Participants

Survey exclusions within the database included individuals living on Indigenous reserves, institutional residences, or serving in the Canadian Armed Forces, collectively representing less than 2.5% of Canada's population (Statistics Canada, 2024). However, the MHACS intentionally oversampled individuals identifying with the four largest visible minority populations in Canada (South Asian, Black, Chinese, Filipino) to address systemic disparities in mental health research and ensure the findings reflect Canada's socio-demographic diversity. To align with the research question's focus on adults aged 20 years or older residing in Canada, participants aged 15-19 years (n = 894) were excluded from the 9861 responses. Missing data for the independent variable, frequency of trouble sleeping, and dependent variables, emotional well-being and positive functionings, were excluded. After applying inclusion criteria and addressing missing data, the final sample comprised 8798 participants.

Study Outcomes

The primary outcome, emotional well-being, was measured using three items from the validated Mental Health Continuum–Short Form (MHC-SF). The MHC-SF was selected for its strong psychometric properties, including high validity and reliability in distinguishing between flourishing and languishing mental health states (Keyes, 2002). Its Likert-scale design enhances sensitivity to subtle variations in well-being, while categorization aligns with clinical frameworks for actionable interpretation (Lamers et al., 2010). Participants rated how frequently they felt happy, interested in life, and satisfied with their life over the past month on a 6-point Likert scale where 0 = Never, 1 = Once or twice, 2 = About once a week, 3 = 2–3 times a week, 4 = Almost every day and 5 = Every day. A composite score classified individuals as having "high emotional well-being" if they reported these positive emotions daily or almost daily for at least one of the three items. The secondary outcome, positive functioning, was assessed using the remaining eleven MHC-SF items, which evaluated psychosocial aspects such as community belonging, self-efficacy, and life purpose. Responses were similarly scored as "high" positive functioning if individuals reported almost every day or every day to six out of eleven items. Missing data was addressed by pairwise deletion for cases with missing sleep disturbance data, and missing data for the dependent variables: emotional well-being and positive functioning.

With a strong internal consistency of $\alpha > 0.80$, indicating that the test is reliable because the score is consistent, and validity in adult populations across diverse cultures including United States, the Netherlands and South Africa, the MHC-SF is a validated tool (Keyes, 2005, 2006, 2009; Keyes et al., 2008; Lamers et al., 2010; Westerhof & Keyes, 2009). MHC-SF's accuracy, reproducibility, validity, and clinical relevance for identifying well-being highlight its appropriateness for this study.

Additional Data Collection

Additional variables (age, sex, visible minority status, and place of residence) were included for the purpose of describing the study population (Table 1). Age was divided into three categories: young adulthood (20-34 years old), middle adulthood (35-64 years old), and later adulthood (65 years old or older) based on developmental psychology frameworks (American

Psychological Association, n.d.). Sex was categorized into Female, Male and not stated. Visible minority status was divided into seven categories including (1) South Asian, (2) Chinese, (3) Black, (4) Filipino, (5) Arab, Latin American, South East Asian, West Asian, Korean, Japanese, and visible minority, (6) not a visible minority and (7) not stated. The not a visible minority category included individuals identifying as White, Indigenous, or those with mixed ethnicities. Place of residence was defined by population size: rural (<1 000 residents), small population centre (1000–29 999), medium population centre (30 000–99 999), and large urban population centre (≥100 000). Residence categories were included to assess potential differences in rural versus urban sleep disturbances and mental health outcomes. Visible minority oversampling ensured equitable representation in line with Canada's public health priorities.

Statistical Analysis

All analyses were conducted using SPSS (v28). Descriptive statistics, including frequencies and percentages, summarized demographic characteristics in Table 1. Given that the dependent variables are nominal and the independent variable is ordinal, cross tabulation tables were utilized to summarize the frequency distributions across categories. These tables provided a clear picture of how different levels of sleep disturbances are associated with variations in emotional well-being and positive functioning and were assisted in translating the data visually into bar charts. Associations between sleep disturbances (independent variable) and both emotional well-being and positive functioning (dependent variables) were evaluated using chi-square tests, with statistical significance set at $p < 0.05$. Chi-square tests are ideally suited for examining relationships between categorical variables, which is what our research question seeks to explore.

Results

There were an original 9861 valid responses to the survey. After excluding participants under the age of 20 years old, 894 cases were deleted. There were 127 missing data cases for our primary dependent variables, high emotional well-being, that were excluded pairwise. There was 1 missing data case for our secondary dependent variable, high positive functioning, that was excluded pairwise. There were 41 missing data cases for our independent variable, frequency of sleep disturbances, that were excluded pairwise. Following the exclusion and inclusion criteria set for individuals in our study as well as pairwise exclusions, a total of 8798 individual cases were observed compared to the original 9861 valid responses.

Table 1 shows the descriptive characteristics of our study population with our participants categorized by the presence (n=6013) or absence (n=2785) of sleep disturbances as depicted by the Mental Health and Access to Care Survey database for 2022. Significance values were excluded from Table 1 because it is meant to only describe participant characteristics for each group rather than test for statistical differences between these characteristics. Majority of our participants fell into the middle adulthood lifestage (35 to 64 years); 44.1% for the sleep disturbances group and 42.9% for the no sleep disturbances group. Responses were fairly even, between male and female participants for both groups (46.8% for the sleep disturbances group and 58.4% for the no sleep disturbances group. Similarly, there was a similar even distribution for visible minority status versus not a visible minority status (39.8% compared to 56.5% for the sleep disturbances group and 51.2% compared to 45%, respectively for the no sleep disturbances group). Place of residence is predominantly large urban population participants (69.1% for the sleep disturbances group and 73.2% for the no sleep disturbances group).

Table 1.

Descriptive characteristics of study population by sleep disturbance status as depicted by the Mental Health and Access to Care Survey database for 2022 (N=8798)

Characteristics		**No Sleep Disturbances (n=2785)**	**Sleep Disturbances (n=6013)**
Categorical variables presented as *n* (%)			
Age (years)			
	Young Adulthood (20 to 34)	676 (24.3)	1660 (27.6)
	Middle Adulthood (35 to 64)	1194 (42.9)	2654 (44.1)
	Later Adulthood (65+)	915 (32.8)	1699 (28.3)
Sex			
	Male	1627 (58.4)	2812 (46.8)
	*Missing	6 (0.2)	5 (0.1)
Visible Minority Status	South Asian	392 (14.1)	477 (7.9)
	Chinese	246 (8.8)	597 (9.9)
	Black	348 (12.5)	513 (8.5)
	Filipino	233 (8.4)	418 (7.0)
	Other visible minority	206 (7.4)	392 (6.5)
	Not a visible minority	1254 (45.0)	3399 (56.5)
	*Missing	106 (3.8)	217 (3.6)

Note. Table 1 presents the descriptive characteristics of the study population (N = 8798), divided into categories based on the presence or absence of sleep disturbances. Categorical variables are expressed as n (%). Percentages may not total 100% due to rounding.

Figure 1 shows the association between frequency of sleep disturbances and high emotional well-being. Data was acquired from a 5x2 cross tabulation table output from the chi-square analysis. There is a downhill slope from the left most category (no sleep disturbances) to the right most category (sleep disturbances all of the time). The sample size for each category from left to right was 2628, 1952, 2129, 775, and 319. The proportions (%) of each group that answered 'yes' to having high emotional well-being were 94.4, 91.6, 87.0, 77.7, and 72.7 from left to right. There was a significant difference (X2 = 348.05; p < 0.001) between all proportions within the analysis.

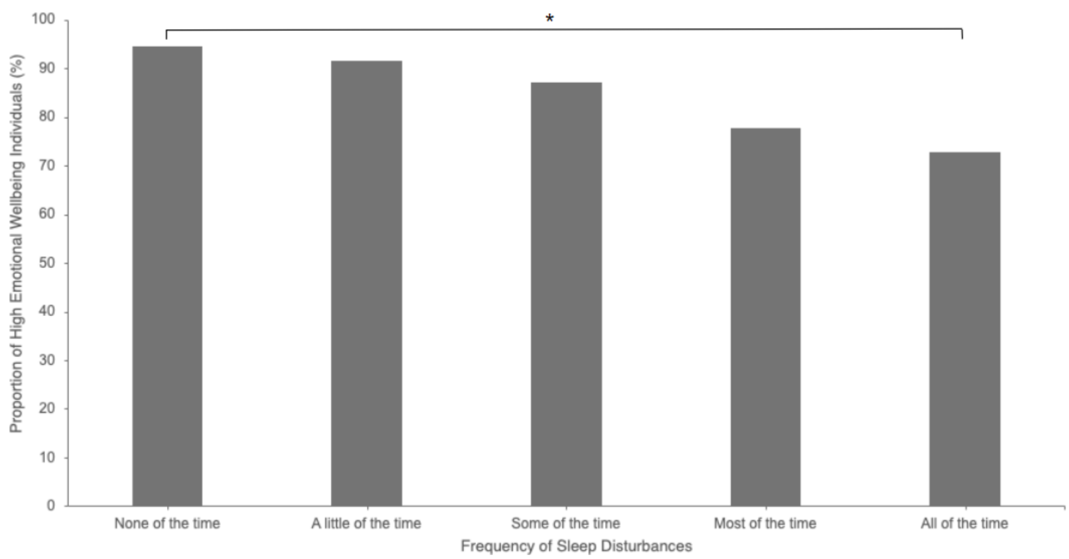


Figure 1. Proportion (%) of participants with high emotional well-being categorized by frequency of sleep disturbances. The pearson chi-square value was X2 = 348.05 and the significance p-value for all groups was <0.001. Significant results are indicated by the asterisk.

Figure 2 shows the association between frequency of sleep disturbances and high positive functioning. Data was acquired from a 5x2 cross tabulation table output from the chi-square analysis. There is a downhill slope from the left most category (no sleep disturbances) to the right most category (sleep disturbances all of the time). The sample size for each category from left to right was 2273, 1511, 1626, 557, and 228. The proportions (%) of each group that answered 'yes' to having high positive functioning were 81.6, 70.9, 66.5, 55.8, and 51.9 from left to right. There was a significant difference (X2 = 360.33; $p < 0.001$) between all proportions within the analysis.

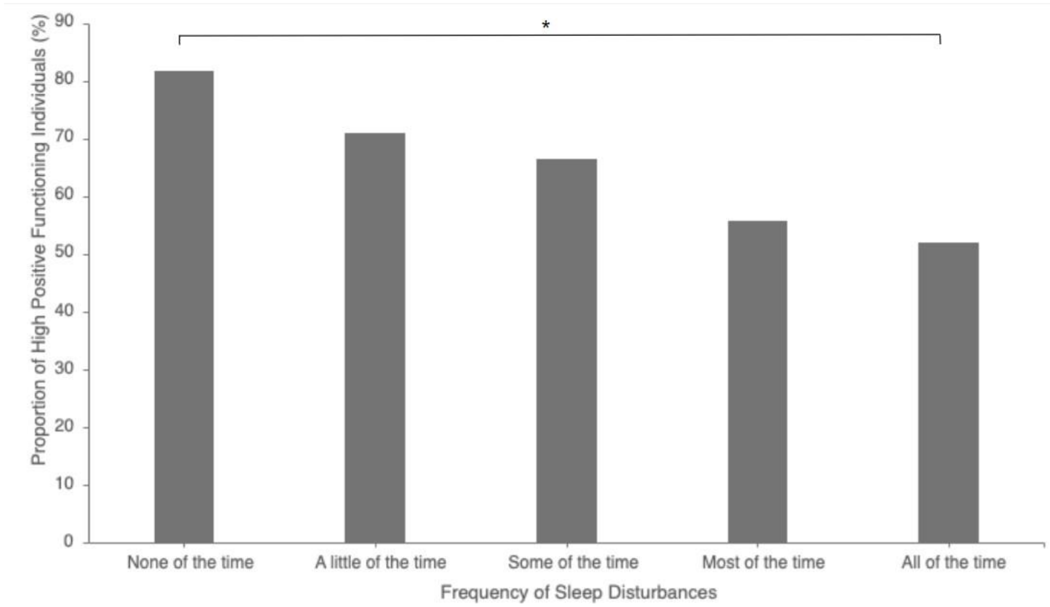


Figure 2. Proportion (%) of participants with high positive functioning categorized by frequency of sleep disturbances. The pearson chi-square value was X2 = 360.33 and the significance p-value for all groups was <0.001. Significant results are indicated by the asterisk.

Discussion

The main findings of the study indicate a significant inverse relationship between self-reported sleep disturbances over the past decade and both emotional well-being and positive functioning among Canadian adults. Participants with more frequent sleep disruptions exhibited lower rates of high emotional well-being and positive functioning. This suggests that Canadian adults who experience sleep challenges, including sleep disorders, frequent sleep disruptions, and inadequate quality or quantity of sleep, may face heightened emotional distress and diminished positive functioning.

Similar to existing literature, our findings align with the hypothesis that sleep quality and quantity are a critical determinant of mental health (Franzen & Buysse, 2008; Goldstein & Walker, 2014; Philbrook & Macdonald-Gagnon, 2021; Steptoe et al., 2008). For instance, Yoo et al. (2007) demonstrated that sleep deprivation impairs the memory encoding of positive emotional stimuli and contributes to a cognitive bias toward negative experiences. This supports our observation of reduced emotional well-being in individuals with frequent sleep disturbances. Similarly, Philbrook and Macdonald-Gagnon (2021) found that poor sleep quality may exacerbate emotional distress and vice versa. While our study is unable to depict the relationship of emotional distress on sleep quality, it does highlight the importance of sleep quality on emotional well-being and positive functioning. This study supports the idea of a bidirectional relationship between sleep quality and psychological well-being seen in our study.

Despite these findings, previous literature has limitations that prevent discoveries from being completely explored. Steptoe et al.'s (2008) study is a cross-sectional design that limits the ability to draw causal conclusions (Steptoe et al., 2008). In addition, their study population is limited to older adults aged 58-72 years old (Steptoe et al., 2008). This reduces the generalizability of their findings to diverse groups. Yoo et al.'s (2007) study looks at single night sleep deprivations, which only provides a snapshot of the full picture and reduces the generalizability for real-world, long-term or chronic impacts of sleep deprivation on the encoding of positive experiences.

Unlike previous studies that are focused on experimental or younger populations, our data provides unique insight into an under-researched scope of sleep-related mental health disparities across diverse demographic groups of Canadian adults aged 20 years and older. In addition, a key difference lies in our inclusion of positive functioning, an overlooked topic in sleep research. Our study revealed that high positive functioning had higher differences in proportions between groups as sleep disturbances increased compared to high emotional well-being. By examining positive functioning alongside emotional well-being, this study explores implications to mitigating psychological suffering and understanding how individuals flourish, aligning with modern models of mental health such as the MHC-SF (Keyes, 2002). According to the MCH model, flourishing mental health is defined as high emotional well-being and high positive functioning. As a result, the study of these variables can influence our understanding of the relationship between sleep disturbances and their impacts on mental health. The oversampling of visible minorities within our study and the MHACS database enhances the generalizability of our findings to Canada's multicultural population, addressing gaps in prior studies that underrepresented these groups.

Like all studies, our study had limitations. Our study is retrospective in nature which can potentially introduce recall bias. Self-reported data in general can be unreliable, especially for subjective measures that are based on the recollection of events occurring over an extended time period (e.g. like sleep disturbances in the past decade). Another limitation is the measurement of variables and the demographic categorization used in the MHACS survey. The "not a visible minority" category is broad and may not reflect the diversity within this group. For instance, Indigenous populations are included in this category, which may obscure important distinctions between caucasian and Indigenous populations. In addition, the MHACS survey categorized 15-19 years of age as one group, despite 18 years of age typically indicating adulthood (Statistics Canada, 2019). As a result, our study excluded this category from our sample population, reducing generalizability to Canadian adults by removing a developmental period that is characterized by heightened sleep and mental health vulnerabilities (Owens et al., 2014). Finally, while the use of a chi-squared analysis allows associations to be identified, the causation or the direction of relationships can not be determined. The chi-squared analysis was also larger than a two by two table making it impossible to tell which proportions are different, or similar, as there is one significance value for the entire table. Despite this limitation, utilizing data from our cross-tabulation table helped create a bar chart that depicted the direction of relationships between variables.

Future studies should build on this study by using prospective designs to track real-time data and reduce recall bias. Objective measures, such as wearable sleep trackers, could complement self-reported data for greater accuracy. Refining demographic categories could also yield more precise insights, particularly regarding Indigenous populations and young adults. The inclusion of continuous variables may also help provide more insight into the direction of relationships by allowing studies to use different types of statistical analyses. Future studies should also explore the role of geographical and socioeconomic disparities on sleep quality and inadvertently, psychological well-being. These confounding variables likely play a role in the experience of sleep, through circadian rhythm misalignment of farmers in rural communities to shift work impacting sleep quantity (Grandner et al., 2017; James et al., 2018). Analyzing these confounding variables will provide implications for public health measures to address disparities that may exist. In light of prior evidence in this field, strengthened by our study, sleep quality plays a critical role in influencing emotional well-being and positive functioning. Thus, addressing sleep disturbances is essential for ensuring flourishing mental health.

References

American Psychological Association. (n.d.). *Adulthood*. American Psychological Association. https://dictionary.apa.org/adulthood

Bibbins-Domingo, K., & Helman, A. (2022). Why Diverse Representation in Clinical Research Matters and the Current State of Representation within the Clinical Research Ecosystem. In www.ncbi.nlm.nih.gov. National Academies Press (US).

Chattu, V. K., Manzar, M. D., Kumary, S., Burman, D., Spence, D. W., & Pandi-Perumal, S. R. (2018). The Global Problem of Insufficient Sleep and Its Serious Public Health Implications. *Healthcare (Basel)*, 7(1), 1-. https://doi.org/10.3390/healthcare7010001

Clayborne, Z. M., Dopko, R. L., Wang, C., Betancourt, M. T., Roberts, K. C., & Capaldi, C. A. (2023). Associations of sleep duration and sleep quality with indicators of mental health among youth and adults: findings from the 2015 Canadian Community Health Survey. *Health Promotion and Chronic Disease Prevention in Canada*, 43(5), 243–259. https://doi.org/10.24095/hpcdp.43.5.04

Clement-Carbonell, V., Portilla-Tamarit, I., Rubio-Aparicio, M., & Madrid-Valero, J. J. (2021). Sleep Quality, Mental and Physical Health: A Differential Relationship. *International Journal of Environmental Research and Public Health*, 18(2), 460-. https://doi.org/10.3390/ijerph18020460

Cross, M. P., Hofschneider, L., Grimm, M., & Pressman, S. D. (2018). Subjective well-being and physical health. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being.* Salt Lake City: DEF Publishers.

Franzen, P.L., & Buysse, D. J. (2008). Sleep disturbances and depression: Risk relationships for subsequent depression and therapeutic implications. *Dialogues in Clinical Neuroscience,* 10(4), 473–481. https://doi.org/10.31887/DCNS.2008.10.4/plfranzen

Gautam, S., Jain, A., Chaudhary, J., Gautam, M., Gaur, M., & Grover, S. (2024). Concept of mental health and mental well-being, it's determinants and coping strategies. Indian *Journal of Psychiatry*, 66(Suppl 2), S231–S244. https://doi.org/10.4103/indianjpsychiatry.indianjpsychiatry_707_23

Goldstein, A. N., & Walker, M. P. (2014). The role of sleep in emotional regulation and stress reactivity. *Annual Review of Clinical Psychology*, 10, 679–702. https://walkerlab.berkeley.edu/reprints/Goldstein-Walker_ARCP-Review_2014.pdf

Grandner, M. A., Williams, N. J., Knutson, K. L., Roberts, D., & Jean-Louis, G. (2016). Sleep disparity, race/ethnicity, and socioeconomic position. *Sleep Medicine*, 18, 7–18. https://doi.org/10.1016/j.sleep.2015.01.020

James, S. M., Honn, K. A., Gaddameedhi, S., & Van Dongen, H. P. A. (2017). Shift Work: Disrupted Circadian Rhythms and Sleep—Implications for Health and Well-being. *Current Sleep Medicine Reports*, 3(2), 104–112. https://doi.org/10.1007/s40675-017-0071-6

Kirkbride, J. B., Anglin, D. M., Colman, I., Dykxhoorn, J., Jones, P. B., Patalay, P., Pitman, A., Soneson, E., Steare, T., Wright, T., & Griffiths, S. L. (2024). The social determinants of mental health and disorder: evidence, prevention and recommendations. *World Psychiatry,* 23(1), 58–90. https://doi.org/10.1002/wps.21160

Keyes, C. L. M. (2002). The Mental Health Continuum: From Languishing to Flourishing in Life. *Journal of Health and Social Behavior*, 43(2), 207–222. https://doi.org/10.2307/3090197

Keyes, Corey L. M. 2005. The subjective well-being of America's youth: Toward a comprehensive assessment. *Adolescent and Family Health*, 4, 3−11.

Keyes, C. L. M. (2006). Mental Health in Adolescence: Is America's Youth Flourishing? American *Journal of Orthopsychiatry*, 76(3), 395–402. https://doi.org/10.1037/0002-9432.76.3.395

Keyes, C. L. M. (2009). Atlanta: *Brief description of the mental health continuum short form (MHC-SF)*. Available from http://www.sociology.emory.edu/ckeyes/

Keyes, C. L. M., Wissing, M., Potgieter, J. P., Temane, M., Kruger, A., & van Rooy, S. (2008). Evaluation of the mental health continuum-short form (MHC-SF) in setswana-speaking South Africans. *Clinical Psychology and Psychotherapy*, 15(3), 181–192. https://doi.org/10.1002/cpp.572

Lamers, S. M. A., Westerhof, G. J., Bohlmeijer, E. T., ten Klooster, P. M., & Keyes, C. L. M. (2011). Evaluating the psychometric properties of the mental health Continuum-Short Form (MHC-SF). *Journal of Clinical Psychology*, 67(1), 99–110. https://doi.org/10.1002/jclp.20741

Kubzansky L. D. , Boehm J. K. , & Segerstrom S. C. (2015). Positive psychological functioning and the biology of health. *Social and Personality Psychology Compass, 9*(12), 645–660.

Kushlev, K., Drummond, D. M., & Diener, E. (2020). Subjective well-being and health behaviors

in 2.5 million Americans. *Appl Psych Health Well-Being, 12*(1), 166–187.

Ngamaba, K. H., Panagioti, M., & Armitage, C. J. (2017). How strongly related are health status and subjective well-being? Systematic review and meta-analysis. *The European Journal of Public Health*, *27*(5), 879–885.

Owens, J. (2014). Insufficient sleep in adolescents and young adults: an update on causes and consequences. *Pediatrics (Evanston)*, 134(3), e921–e932. https://doi.org/10.1542/peds.2014-1696

Paller, K. A., Creery, J. D., & Schechtman, E. (2021). Memory and Sleep: How Sleep Cognition Can Change the Waking Mind for the Better. *Annual Review of Psychology*, 72(1), 123–150. https://doi.org/10.1146/annurev-psych-010419-050815

Philbrook, L. E., & Macdonald-Gagnon, G. (2021). Bidirectional relations between sleep and emotional distress in college students: Loneliness as a moderator. *The Journal of Genetic Psychology*, 182(5), 361–373. https://doi.org/10.1080/00221325.2021.1913982

Sin, N. L., Almeida, D. M., Crain, T. L., Kossek, E. E., Berkman, L. F., & Buxton, O. M. (2017). Bidirectional, Temporal Associations of Sleep with Positive Events, Affect, and Stressors in Daily Life Across a Week. *Annals of Behavioral Medicine*, 51(3), 402–415. https://doi.org/10.1007/s12160-016-9864-y

Statistics Canada. (2024). *Mental Health and Access to Care Survey [Canada]*. Government of Canada. https://doi.org/10.5683/SP3/ABG9CY, Borealis, V5, UNF:6:jz47ZqEtwUOoFN4NwHGdEA

Statistics Canada. (2019, May 14). *Classification of adult status of person.* Government of Canada. https://www23.statcan.gc.ca/imdb/p3VD.pl?Function=getVD&TVD=490890&CVD=490890&CLV=0&MLV=1&D=1

Steptoe, A., O'Donnell, K., Marmot, M., & Wardle, J. (2008). Positive affect, psychological well-being, and good sleep. *Journal of Psychosomatic Research*, 64(4), 409–415. https://doi.org/10.1016/j.jpsychores.2007.11.008

Tout, A. F., Jessop, D. C., & Miles, E. (2024). Positive psychological traits predict future sleep quality and quantity: exploring emotion regulation as a common mediator. *Psychology & Health*, 1–23. https://doi.org/10.1080/08870446.2024.2314722

Walker, M. P., & van der Helm, E. (2009). Overnight Therapy? The Role of Sleep in Emotional Brain Processing. *Psychological Bulletin*, 135(5), 731–748. https://doi.org/10.1037/a0016570

Westerhof, G. J., & Keyes, C. L. M. (2010). Mental Illness and Mental Health: The Two Continua Model Across the Lifespan. *Journal of Adult Development*, 17(2), 110–119. https://doi.org/10.1007/s10804-009-9082-y

Yoo, S. S., Hu, P. T., Gujar, N., Jolesz, F. A., & Walker, M. P. (2007). A deficit in the ability to form new human memories without sleep. *Nature Neuroscience*, 10(3), 385–392. https://doi.org/10.1038/nn1851

The Association of Healthcare Access Factors in Delays on Insulin Treatment in Canadian Adults: A Cross-Sectional Study

Zainab Amer and Nadia Jaipersaud

McMaster University

Department of Interdisciplinary Science

LIFESCI 3LX3: Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8, 2025

The Association of Healthcare Access Factors in Delays on Insulin Treatment in Canadian Adults: A Cross-Sectional Study

Introduction

Diabetes affects over 3.7 million Canadians, and this number continues to rise each year (Government of Canada, 2023). This chronic condition results in elevated blood glucose levels due to the body's inability to produce enough insulin or effectively use the insulin it makes (World Health Organization, 2023). Although insulin therapy plays a critical role in diabetes management, many individuals do not receive it in a timely manner (World Health Organization, 2023). El-Sayed Moustafa et al. (2023) emphasizes that therapeutic inertia, the failure to initiate or intensify insulin therapy, leads to serious complications like heart disease and kidney damage. They report that delays often stem from patient fears, provider hesitation, and fragmented care systems. These delays increase the risk of heart failure by 64%, myocardial infarction by 67%, and stroke by 51%. Khunti et al. (2013) analyzed data from over 80,000 individuals with type 2 diabetes and found that many patients experience a median delay of 1.6 years before their insulin treatment was initiated.

Available literature indicates that access to healthcare services significantly impacts insulin initiation. Harris et al. (2010) investigated this relationship in Canadian family practice settings and found that patients waited an average of 9.2 years after diagnosis to begin insulin. Their findings suggest that primary care provider behavior, confidence, and follow-up practices strongly influence whether and when patients start insulin. The study also highlighted how the lack of integrated diabetes support, such as nurse educators and structured insulin protocols further complicates timely treatment. Tran et al. (2024) reinforced these concerns in a retrospective cohort study in Alberta. They tracked individuals with type 2 diabetes over an eight-year period and found that fewer than half of eligible patients initiated basal insulin despite meeting clinical criteria. This pattern points to barriers in the primary care system, including time constraints, limited clinical support, and insufficient access to diabetes educators.

Healthcare insurance also plays a major role in insulin initiation. Booth et al. (2021) conducted a national population-based study and found that insulin pump use was significantly higher in provinces offering government funding (47.8%) than in those without it (37.7%). Even after adjusting for demographics and geography, patients in funded provinces had 1.45 times higher odds of using insulin pumps. A comprehensive report by Diabetes Canada and the Library of Parliament (n.d) reveals that individuals without private insurance bear significant out-of-pocket costs up to $2,600 annually for injections and supplies, and as much as $4,900 per year for insulin pump therapy. Although provincial healthcare plans cover physician visits and some drug costs, they often do not cover essential supplies like glucometers, test strips, or insulin delivery devices, creating financial barriers that delay or discourage timely insulin initiation.

Research indicates that self-perceived unmet healthcare also plays an important role in insulin initiation. According to Sanmartin et al. (2006), approximately 13% of Canadians aged 15 and older reported at least one instance in the past year where they felt they needed healthcare but did not receive it. Individuals with lower income, poorer self-rated health, and no regular doctor were significantly more likely to report unmet needs. These structural and perceived barriers may have downstream effects on diabetes care, particularly on the timely initiation of insulin therapy, which often requires regular follow-up, coordinated care, and access to diabetes education and supplies.

Employment status and personal income also influence diabetes care and may contribute to delays in insulin initiation. Several Canadian studies have identified income-related disparities in diabetes outcomes. For example, Booth et al. (2012) found that individuals living in lower-income neighborhoods were less likely to receive recommended diabetes treatments, including insulin, and were more likely to experience complications. Glazier et al. (2011) noted that job instability and lack of workplace benefits can negatively affect disease management, particularly among those without extended health coverage. Although these socioeconomic factors are not always the primary focus of insulin initiation research, they represent important covariates that may intersect with other access barriers to influence treatment decision

Despite growing awareness of the risks associated with delayed insulin initiation and the role of healthcare access in diabetes management, few studies have examined how specific access-related factors interact to influence the timing of insulin therapy. Research has individually linked primary care access, health insurance coverage, and socioeconomic disadvantage to insulin delays, yet these factors are often studied in isolation. In particular, there is limited Canadian literature that investigates how self-perceived unmet healthcare needs, a subjective but powerful measure of healthcare access, relates to the time between diabetes diagnosis and the start of insulin therapy. While studies acknowledge that patients with fewer resources may delay treatment, researchers have not fully explored how perceived gaps in care contribute to therapeutic inertia. This lack of integrated analysis creates an important gap in the literature, especially in the Canadian context where public and private healthcare coverage intersect. To address this gap, the study primarily investigates the question: Is an association between access to a regular healthcare provider and health insurance with the time between diagnosis and start of insulin therapy among Canadians aged 18 and older? The secondary aim is to determine if there is a relationship between self-perceived unmet healthcare needs and the time between diagnosis and start of insulin therapy among Canadians aged 18 and older.

Methods

Study Design and Database

This cross-sectional observational study used data from the Canadian Community Health Survey (CCHS) 2017–2018 annual component. Statistics Canada developed and administers the CCHS to collect nationwide information on health status, healthcare access, and social determinants of health among Canadians aged 12 and older. The survey applies a multistage stratified cluster sampling method to ensure representation across regions and demographic groups. Data collection occurred through computer-assisted interviews, conducted either in person or by telephone. Provincial and territorial governments selected optional modules to address region-specific priorities, allowing the survey to capture diverse health-related information. The CCHS was selected for its comprehensive, population-level data relevant to healthcare access and chronic disease management.

Study Participants

The dataset contained information collected from 9292 participants from the CCHS. There were specific inclusion criteria utilized while selecting participants. Only Canadians over the age of 18 with a diabetes diagnosis that were questioned about their delay in insulin therapy were included. Considering that the predictor variables analyzed in this data set are access to a regular healthcare providers, health insurance, and self-perceived unmet needs, the adult age group was selected for increased accuracy in answers. Adults are more likely to engage independently with the healthcare system and make informed decisions regarding insurance coverage and treatment options, making them more appropriate for analyses of healthcare access and perceived barriers (Bailey et al., 2017). We investigated participants who had delays in the initiation of insulin therapy and used those who had never started insulin therapy as a baseline. Participants who responded with “valid skip”, “don’t know”, “refusal”, or “not stated” were excluded.

Study Outcomes

The primary outcome was the delay of the start of insulin therapy from the point of the diabetes diagnosis. The start of insulin therapy measured from the point of diabetes diagnosis is an ordinal categorical variable that was self-reported. The variable is measured through the answer choices of ≤ 2 months, 2 months to ≤ 6 months, 6 months to ≤ 1 year, ≥ 1 year, and never. The variables were collapsed to these final values to accurately define the association of access to a healthcare provider and insurance on the delay of beginning insulin therapy. These predictors were distributed in a dichotomous categorical manner, only data from the values “yes” and “no” were deemed relevant to reflect the known-system level barriers that diabetes patient face. According to an article by Yochioka et al. (2014), patients often lack a proper understanding of their illness and the role of insulin, causing delays in initiation.

The primary outcome of the secondary question also focused on the time between diagnosis and the start of insulin treatment as well. Participants were questioned on their perception of their unmet needs during the same time period as they were waiting for the start of their insulin therapy. In addition to systemic behaviors, a patients’ beliefs often has an impact on the initiation of insulin therapy as well. Clinical studies have identified attitudinal behaviours and negative self-perceptions as reasons for delaying insulin therapy (Brod et al., 2009). In the database, this was measured by using the “self-perceived unmet needs” inclusion category. While participants were asked further questions on the location of where they felt that their needs were unmet, this data was not deemed relevant. The answer choices analyzed were “yes” and “no”

Additional Data Collection

`Additional data included in Table 1 from the database are sex, province, age, personal income, and employment. Personal income and employment are important confounding factors that are covariates when analyzing the participants' access to a regular healthcare provider and insurance. These variables were also used to generate Figure 2, which visually displays their distribution in relation to insulin initiation. Although sex and province were not directly used in the analyses of this report, they serve to describe the population and may require additional investigation to sufficiently fill the knowledge gap in future studies.

Statistical Analysis

IBM SPSS software was utilized to run the statistical analysis through the manipulation of the original 2017-2018 CCH dataset found on ODESI. Table 1 summarizes the frequency and percentage of individuals who fell into categories based on sex, age, province, income, employment, healthcare insurance, access to regular healthcare provider, and self-perceived unmet needs. All variables used in the analysis were categorical variables, so a normality test was not required. The variables measuring access to regular healthcare, healthcare insurance and self perceived unmet needs were used as predictors in determining their association with the dependent variable, which measured the time between diagnosis and the start of insulin therapy.

A multinomial logistic regression analysis was performed to determine whether the predictor variables were significantly related to the delay in insulin therapy. The p-value of less than 0.05 was used as a marker of statistical significance for the multinomial logistic regression.

Results

The 2017-2018 Canadian Community Health Survey included responses of 10 053 participants that were over the age of 18 and diagnosed with diabetes. Exclusion and inclusion criteria was applied, where missing data was removed based on participants who answered "refusal", "not stated", "don't know" for the dependent variable, time between diabetes diagnosis and start of insulin treatment. For this study, 9292 (92%) of these participants were used to analyze the association between the predictor variables and the dependent variable. The descriptive characteristics of the population demographics is displayed in Table 1.

Table 1.

Descriptive Characteristics of the Study Population (n=9292).

Characteristic		Time between diabetes diagnosis and start of insulin treatment					p-value
		<2 months n=925	2 months to <6 months (n=89)	6 months to <1 year (n=109)	1 year or more (n=1702	Never (baseline) (n=6467)	
Sex (N = 9292)							0.006
	Male	514 (55.57)	38 (42.70)	43 (39.45)	892 (52.41)	3439 (53.18)	
	Female	411 (44.43)	51 (57.30)	66 (60.55)	810 (47.60)	3028 (46.82)	
Age (N=9292)							
	18+	925 (100)	89 (100)	109 (100)	1702 (100)	6467 (100)	
Province (N=9292)							<0.001
	Newfoundland and Labrador	38 (4.11)	4 (4.50)	4 (3.67)	70 (4.11)	234 (3.62)	
	PEI	36 (3.89)	3 (3.37)	2 (1.83)	50 (2.94)	115 (1.78)	
	Nova Scotia	47 (5.08)	2 (2.25)	8 (7.34)	113 (6.64)	319 (4.93)	
	New Brunswick	32 (3.50)	4 (4.50)	7 (6.42)	79 (4.64)	267 (4.13)	
	Quebec	151 (16.32)	23 (25.84)	16 (14.68)	265 (15.57)	1449 (22.41)	
	Ontario	294 (31.78)	25 (28.09)	42 (38.53)	586 (34.43)	2053 (31.75)	
	Manitoba	40 (4.32)	4 (4.50)	5 (4.59)	83 (4.88)	301 (4.65)	
	Saskatchewan	47 (5.08)	4 (4.50)	6 (5.50)	81 (4.76)	272 (4.21)	
	Alberta	118 (12.76)	4 (4.50)	7 (6.42)	161 (9.46)	646 (9.99)	
	British Columbia	105 (11.35)	14 (15.73)	11 (10.09)	175 (10.28)	718 (11.10)	
	Yukon	8 (0.86)	1 (1.12)	1 (0.92)	16 (0.94)	39 (0.60)	
	Northwest Territories	7 (0.76)	1 (1.12)	0 (0.00)	19 (1.12)	37 (0.57)	
	Nunavut	2 (0.22)	0 (0)	0 (0.00)	4 (0.24)	17 (0.26)	
Personal Income (N=9278)							0.001
	No Income or Income Loss	8 (0.87)	2 (2.25)	3 (2.75)	14 (0.82)	66 (1.02)	
	Less than $20,000	298 (32.29)	33 (37.08)	36 (33.03)	548 (32.24)	1834 (28.40)	
	$20,000 to $39,999	312 (33.80)	36 (40.45)	35 (32.11)	634 (37.29)	2286 (35.40)	
	$40,000 to $59,999	140 (15.17)	11 (12.36)	20 (18.35)	253 (14.88)	1081 (16.74)	
	$60,000 to $79,999	68 (7.37)	3 (3.37)	7 (6.42)	116 (6.82)	576 (8.92)	
	$80,000 or More	97 (10.51)	4 (4.49)	8 (7.34)	135 (7.94)	614 (9.51)	
Employment (N = 2786)							0.069
	Employee	317 (85.91)	21 (95.45)	23 (88.46)	327 (80.54)	1593 (81.15)	
	Self-employed	52 (14.09)	1 (4.55)	3 (11.54)	79 (19.46)	370 (18.85)	
Healthcare Insurance (N=9292)							0.981
	Yes	40 (4.32)	4 (4.49)	5 (4.59)	83 (4.88)	301 (4.65)	
	No	885 (95.68)	85 (95.51)	104 (95.41)	1619 (95.12)	6166 (95.35)	
Regular Healthcare Provider (N=9292)							0.379
	Yes	72 (94.47)	86 (97.73)	103 (96.26)	1615 (95.28)	6082 (94.43	
	No	51 (5.53)	2 (2.27)	4 (3.74)	80 (4.72)	359 (5.57)	
Self-Perceived Unmet Needs (N=9292)							0.894
	Yes	533 (57.62)	49 (55.06)	65 (59.63)	959 (56.35)	3641 (56.30)	
	No	392 (42.38)	40 (44.94)	44 (40.37)	743 (43.65)	2826 (43.70)	

Note. Categorical variables are presented as n (column %). The reference group comprises patients who never initiated insulin treatment. Primary analysis includes N=9,292 observations, except employment status (subgroup N=2,786). All percentages reflect complete cases analysis only. HC = healthcare.

A multinomial logistic regression was carried out to assess the association between the predictor variables of access to a regular healthcare provider, health insurance, and self-perceived unmet needs on the delay in the initiation of insulin therapy for Canadian adults diagnosed with diabetes. According to Table 2, the coefficient ratio, odds ratio and significance were displayed for all 3 predictor values through the B, Exp(B), and Wald columns respectively. As illustrated in the table, no associations were considered statistically significant.

Table 2.

Multinomial Logistic Regression Output Displaying Association Between Predictor Variables and Time to Initiation of Insulin Therapy (n=9292).

	B			Sig.			Exp(B)			Wald			df		
	HC_Provider	Insurance	Unmet_Need	HC_Provider	Insurance	Unmet_Need	HC_Provider	Insurance	Unmet_Need	HC_Provider	Insurance	Unmet_Need	HC_Provider	Insurance	Unmet_Need
Less than 2 months	0.11	-0.077	0.05	0.944	0.655	0.478	1.011	0.926	1.052	0.005	0.2	0.503	1	1	1
2 months to less than 6 months	0.931	-0.033	-0.021	0.194	0.949	0.921	2.537	0.967	0.979	1.685	0.004	0.01	1	1	1
6 months to less than 1 year	0.421	0.002	0.108	0.411	0.996	0.585	1.524	1.002	1.114	0.675	0	0.298	1	1	1
1 year or more	0.175	0.036	0.004	0.167	0.78	0.948	1.192	1.036	1.004	1.91	0.078	0.004	1	1	1

Multinomial logistic regression results comparing time to insulin initiation across delay categories (<2 months, 2–6 months, 6 months to less than <1year, ≥1 year) against the reference group ('Never'). Coefficients (B), odds ratios [Exp(B)], and p-values (Sig.) are reported for three predictors: healthcare provider access (HC_Provider), health insurance coverage (Insurance), and self-perceived unmet needs (Unmet_Need). No associations reached statistical significance (all p > 0.05), though healthcare provider access showed marginally elevated odds of delay in the 2–6 month category (OR = 2.54, p = 0.194). Wald statistics indicate weak evidence for individual predictor effects.

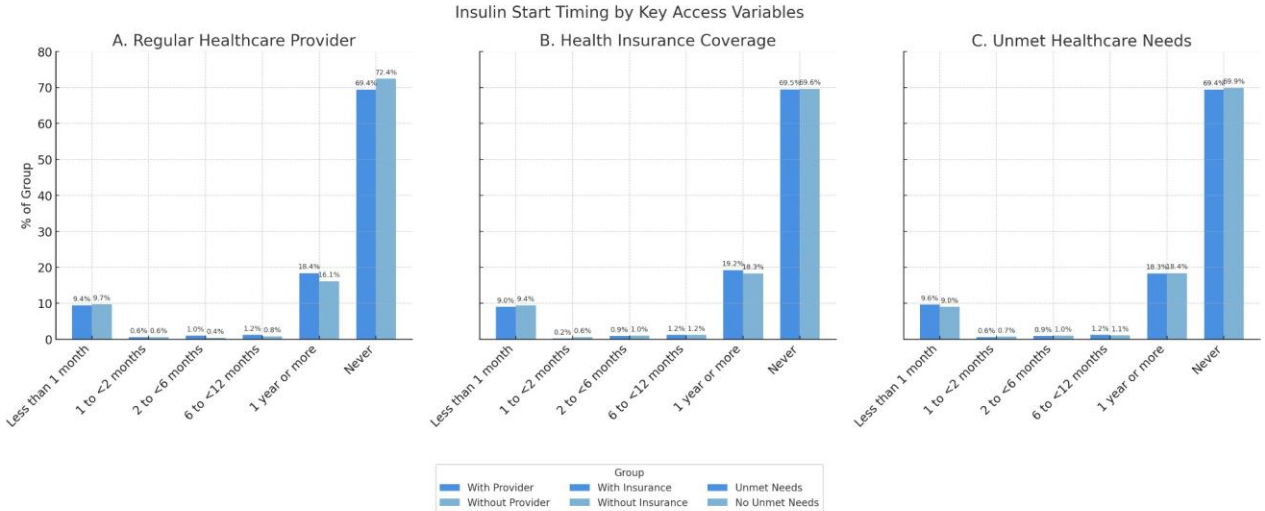


Figure 1.0. Insulin Start Timing by Key Access Variables (n=9292). Three clustered bar charts displayed the percentage of individuals in each insulin initiation timing category, grouped by (A) regular healthcare provider, (B) healthcare insurance coverage, and (C) unmet healthcare needs. The graphs used data from the 2017–2018 Canadian Community Health Survey (CCHS) and included only adults aged 18 and older with a self-reported diabetes diagnosis. "Never started insulin" served as the baseline category for comparison. Small differences were observed, such as a slightly higher proportion of insulin non-initiation among participants without a regular provider (72.4%) compared to those with one (69.4%). However, no substantial differences were found between groups overall. These results suggested that, within this sample, access-related factors had a limited impact on the timing of insulin initiation.

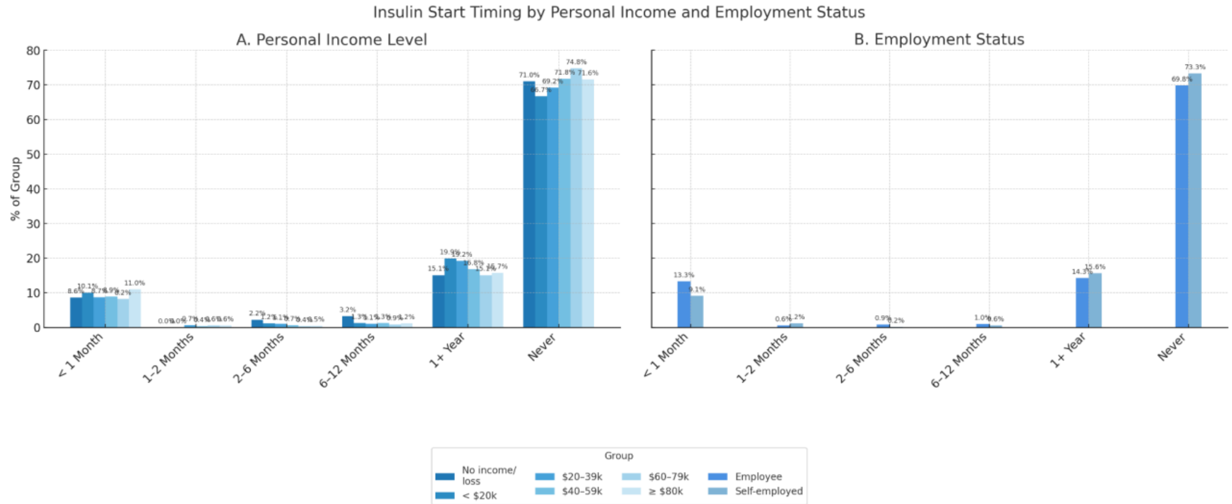


Figure 2.0. Insulin Start Timing by Personal Income and Employment Status (n = 9292). Two clustered bar charts showed the percentage of individuals in each insulin initiation timing category, grouped by (A) personal income level and (B) employment status. Data from the 2017 to 2018 Canadian Community Health Survey (CCHS) were used. The results were limited to adults aged 18 and older with a self-reported diabetes diagnosis. Across all income and employment groups, the majority of participants reported never starting insulin, with relatively small differences observed between groups. "Never started insulin" served as the reference category. Compared to this group, participants in the lowest (< $20,000) and highest (≥ $80,000) income brackets had slightly higher rates of early insulin initiation (< 1 month), while middle-income groups showed lower rates. Moderate delay categories (1–12 months) remained low and consistent across all groups. These results suggested no clear pattern of delayed insulin use among lower-income individuals in this sample.

Discussion

This study aimed to primarily investigate whether various healthcare factors such as access to regular healthcare providers and insurance impacted the time between diagnosis and start of insulin therapy. The study explored whether self-perceived unmet healthcare needs are associated with the delay in insulin therapy to provide an analysis of the social determinants at play. The target group was Canadians over the age of 18 with a confirmed diabetes diagnosis for both of these outcomes. While none of the analyses yielded a result with statistical significance, there is a notable difference seen in the logistic regression model. Within the 2 to 6 months category, the odds ratio of OR=2.54 signifies a 2.5x higher chance of participants receiving treatment with a regular healthcare provider. This displays that participants have a higher chance of receiving earlier treatment if they have a healthcare provider.

As indicated in Figure 1, across the entire sample (N = 9,252), 69.6% of participants never started insulin after a diabetes diagnosis. This was consistent across nearly all subgroups, suggesting a widespread delay in insulin initiation among Canadian adults with diabetes. Both the lowest (<$20k) and highest (≥$80k) income groups had slightly higher rates of early insulin initiation (<1 month) compared to middle-income groups. Self-employed individuals were slightly more likely to never start insulin (73.3%) compared to employees (69.8%), while employees were more likely to initiate insulin within 1 month of diagnosis (13.3% vs. 9.1%). However, these differences were small and not statistically significant.

Our findings provide a nuanced contribution to the growing Canadian literature on insulin initiation and healthcare access. Previous studies have found that access-related barriers often delay treatment for individuals with diabetes (Canadian Institute for Health Information, 2023). Although our logistic regression identified increased odds of insulin initiation within 2–6 months for participants with a regular provider (OR = 2.54), this result was not statistically significant and may be due to sparse data in that subgroup. Harris et al. (2010) found that in Canadian family practice settings, patients waited an average of 9.2 years to begin insulin after diagnosis. The research emphasized the role of physician behaviour and follow up practices as contributors to receiving timely insulin treatment. Kiran et al. (2016) found that availability of primary care physicians was associated with higher rates of adherence to insulin initiation. However, our findings suggest that having access to a regular healthcare provider was not significantly associated with earlier insulin initiation in our sample.

In examining the role of healthcare insurance, our findings revealed no significant difference in insulin initiation timing between insured and uninsured participants. Our data showed that approximately 69.5% of insured and 69.6% of uninsured participants never initiated insulin, suggesting that insurance status alone may not explain delays in treatment initiation. This is somewhat inconsistent with previous Canadian research, such as studies by Booth et al. (2012), which found that lack of drug coverage was associated with poorer diabetes outcomes and reduced medication adherence. However, it is worth noting that insulin itself, along with related supplies (e.g., needles, glucose monitors), is not always fully covered, particularly for adults under 65 without private or employer-provided insurance.

We investigated whether self-perceived unmet healthcare needs influenced the timing of insulin initiation among Canadian adults with diabetes. Our analysis revealed no significant association between unmet needs and insulin initiation, approximately 69.4% of participants reported unmet needs and 69.9% of those without unmet needs never started insulin therapy. This finding suggests that self-perceived barriers to healthcare access may not directly impact insulin initiation decisions.This result contrasts with previous Canadian research indicating that unmet healthcare needs are linked to poorer health outcomes and delayed treatment. For instance, Allin, Grignon, and Le Grand (2010) found that individuals reporting unmet needs due to access barriers did not consistently show patterns of healthcare utilization, suggesting potential equity implications.

The analysis of income and insulin initiation revealed no clear linear association. Participants in both the lowest (< \$20,000) and highest (≥ \$80,000) income groups initiated insulin slightly earlier than those in middle-income groups. For example, 10.1% of participants in the lowest income group started insulin within one month, compared to 8.2% in the \$60,000–79,999 group. This nonlinear pattern challenges the assumption that lower income consistently leads to treatment delays, contrasting with findings from several Canadian studies. A population-based analysis by Booth et al. (2012) reported that individuals residing in lower-income neighborhoods were less likely to receive appropriate diabetes medications, including insulin, suggesting that economic disadvantage frequently impairs timely care.

Employment status analysis revealed only minor differences between groups. Self-employed individuals were slightly more likely to never initiate insulin (73.3%) compared to employees (69.8%), while employees were more likely to begin insulin within one month (13.3% vs. 9.1%). Although these differences were not statistically significant, they may reflect disparities in health insurance coverage, time flexibility, or healthcare-seeking behavior. Canadian studies such as Glazier et al. (2011) have shown that job-related instability and lack of workplace benefits can negatively impact chronic disease management. These findings suggest that while employment type may influence insulin use, it is likely intertwined with other social determinants such as insurance status, access to extended drug plans, or competing work demands.

This study presents several limitations that should be considered when interpreting the findings. The analysis relies on self-reported data from the Canadian Community Health Survey, which introduces the possibility of recall bias and misclassification, particularly for variables such as insulin initiation timing and self-perceived unmet healthcare needs. The cross-sectional design also restricts the ability to draw causal inferences or track changes in insulin use over time. In addition, the outcome variable for insulin initiation was divided into broad time intervals, which may have obscured more nuanced patterns and led to unstable regression estimates due to sparse data in less-populated categories. While the study included key access and socioeconomic factors, it did not account for other important social determinants of health such as education, race and ethnicity, housing insecurity, or health literacy. These unmeasured variables may play a significant role in shaping both healthcare access and treatment decisions.

This study found no significant associations between insulin initiation timing and access-related variables such as having a regular healthcare provider, health insurance coverage, or income level. These findings suggest that commonly measured access indicators may not fully explain why so many individuals delay or avoid insulin therapy. To address this gap, future research should explore less visible influences, including patient perceptions of insulin, communication dynamics with healthcare providers, and therapeutic inertia. Longitudinal studies can offer deeper insight into how these factors develop and interact over the course of care, while qualitative research could reveal emotional or psychological barriers, such as fear of injections or concerns about dependency. More broadly, these findings highlight the need for public health strategies that extend beyond access improvements. Strengthening diabetes education, addressing stigma, and equipping primary care providers to initiate timely insulin therapy may be essential to improving diabetes outcomes in Canada.

References

Allin, S., Grignon, M., & Le Grand, J. (2010). Subjective unmet need and utilization of health care services in Canada: What are the equity implications? *Social Science & Medicine*, 70(3), 465–472. https://doi.org/10.1016/j.socscimed.2009.10.027

Booth, G. L., Bishara, P., Lipscombe, L. L., Shah, B. R., Feig, D. S., & Bierman, A. S. (2012). Universal drug coverage and socioeconomic disparities in major diabetes outcomes. *Diabetes Care*, 35(11), 2257–2264. https://doi.org/10.2337/dc12-0260

Booth, G. L., Kapral, M. K., Fung, K., & Tu, J. V. (2021). Relation between age and use of insulin pumps in adults with type 1 diabetes. *CMAJ*, 193(21), E743–E750. https://doi.org/10.1503/cmaj.200973

Brod, M., Kongsø, J. H., Lessard, S., & Christensen, T. L. (2009). Psychological insulin resistance: Patient beliefs and implications for diabetes management. *Quality of Life Research*, 18(1), 23–32. https://doi.org/10.1007/s11136-008-9419-1

Canadian Institute for Health Information. (2023). *Access to primary care: Many Canadians face challenges*. https://www.cihi.ca/en/primary-and-virtual-care-access-emergency-department-visits-for-primary-care-conditions/access-to-primary-care-many-canadians-face-challenges

Diabetes Canada & Library of Parliament. (n.d.). *Out-of-pocket costs for insulin therapy and supplies in Canada*. https://www.diabetes.ca/advocacy/reports/out-of-pocket-costs-report

El-Sayed Moustafa, J. S., Faruque, L. I., & Tang, T. S. (2023). Therapeutic inertia in diabetes care: An evolving global challenge. *Current Diabetes Reports*, 23, 129–140. https://doi.org/10.1007/s11892-023-01554-y

Glazier, R. H., Gozdyra, P., & Yeritsyan, N. (2011). *Geographic variation in primary care need, service use and provider availability in Ontario: An ICES atlas*. Institute for Clinical Evaluative Sciences.

Government of Canada. (2023). *Diabetes: Overview*. https://www.canada.ca/en/public-health/services/chronic-diseases/diabetes.html

Harris, S. B., Kapor, J., Lank, C. N., & Willan, A. R. (2010). Clinical inertia in patients with T2DM requiring insulin in family practice. *Canadian Family Physician*, 56(12), e418–e424. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3001537/

Khunti, K., Davies, M. J., & Kosiborod, M. N. (2013). Clinical inertia in people with type 2 diabetes: A retrospective cohort study of more than 80,000 individuals. *Diabetes Care*, 36(11), 3411–3417. https://doi.org/10.2337/dc13-0331

Kiran, T., Kopp, A., Moineddin, R., Victor, J. C., Campbell, R. J., & Glazier, R. H. (2016). Impact of primary care model type on specialist referral rates in Ontario, Canada: A cross-sectional study. *CMAJ Open*, 3(1), E37–E44. https://doi.org/10.9778/cmajo.20140073

Sanmartin, C., Houle, C., Tremblay, S., & Berthelot, J.-M. (2006). Access to health care services in Canada, 2003. *Health Reports*, 17(1), 21–31. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1551962/

Starfield, B., Shi, L., & Macinko, J. (2005). Contribution of primary care to health systems and health. *The Milbank Quarterly*, 83(3), 457–502. https://www.milbank.org/quarterly/articles/contribution-of-primary-care-to-health-systems-and-health/

Tran, L. M., Ronksley, P. E., & Quan, H. (2024). Delays in insulin initiation among Canadian adults with type 2 diabetes: A retrospective cohort study. *Canadian Journal of Diabetes*, 48(1), 11–18. https://doi.org/10.1016/j.jcjd.2023.08.004

World Health Organization. (2023, April 6). *Diabetes*. https://www.who.int/news-room/fact-sheets/detail/diabetes

Yoshioka, N., Ishii, H., Tajima, N., Iwamoto, Y., & DAWN Japan group. (2014). Differences in physician and patient perceptions about insulin therapy for management of type 2 diabetes: The DAWN Japan study. *Current Medical Research and Opinion*, 30(2), 177–183. https://doi.org/10.1185/03007995.2013.855187

Nutrition, Income, and Perceived Health: Examining Fruit and Vegetable Consumption in Canadian Adults with Musculoskeletal Conditions

Emma Chittick, Rebecca Purkis, Yana Vaynberg

McMaster University

School of Interdisciplinary Sciences

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8th, 2025

Nutrition, Income, and Perceived Health: Examining Fruit and Vegetable Consumption in Canadian Adults with Musculoskeletal Conditions

Introduction

Chronic musculoskeletal disorders (CMDs), including osteoporosis, arthritis, and fibromyalgia, constitute a leading cause of disability globally and impose financial burdens on healthcare systems (Reginster, 2002). Osteoporosis is characterized by a change in bone microarchitecture, loss of bone mass, and reduced mineral density, increasing fracture risk (Rachner et al., 2011). Arthritis manifests as joint stiffness and tenderness from inflammation, hindering mobility over time, while fibromyalgia causes musculoskeletal tenderness from inflammation (Krsnich-Shriwise, 1997; Reginster, 2002). Typically, the onset age of CMDs is around 50 years old, except for fibromyalgia, which has an earlier onset period that can emerge around 35 years old (Rachner et al., 2011). With older adults being more susceptible to CMDs and representing the fastest growing cohort, addressing CMDs progression and impact is crucial (Nicklett & Kadell, 2013).

Diet is a key modifiable factor in CMD prevention. Around 70-90% of chronic diseases are preventable through modified lifestyle-related factors (Wagenaar et al., 2021). Increased fruit and vegetable intake in older adults reduces the risk of developing CMDs, reduces inflammation, and improves musculoskeletal function (Nicklett & Kadell, 2013). However, the relationship between fruit and vegetable consumption, explicitly looking at consumption dosage and self-reported perceived health in individuals with CMDs, remains understudied, particularly in Canada's adult population (Nisa & Kurotani, 2023). Additionally, household income influences dietary choices as they face financial barriers to accessing nutritious food, potentially exacerbating health disparities (Nisa & Kurotani, 2023). Examining how household income impacts fruit and vegetable consumption can provide further insight into dietary inequities and their consequences (Reginster, 2002).

Dietary trends amongst older adults warrant concern as nutrient-rich fruits and vegetables, like dark green and orange vegetables, are only 12 -15% of total vegetable intake

(Nicklett & Kadell, 2013). Another study by Nicklett and Kadell (2013) found that in adults with CMDs, diets rich in nutrient-dense fruits and vegetables high in calcium and vitamin D were associated with improved bone mineral density. This is particularly relevant to our study as maintaining bone health is essential for managing progression of CMDs. Supporting this, Pena et al. (2020) reveal 96% of participants with musculoskeletal pain had vitamin D deficiency. Other nutrients, like magnesium, potassium, vitamin C, and vitamin K, are also high in fruits and vegetables, preventing bone loss (Nicklett & Kadell, 2013). A three-year longitudinal follow-up study by Lewis et al. (2019) with nearly 400 adults found diets high in potassium reduced muscle loss in adults over 65 years old. Collectively, these findings emphasize that adequate consumption of nutrient-rich fruits and vegetables influences both physical and perceived health of adults with CMDs. Despite benefits, many individuals fall short of the recommended daily fruit and vegetable intake (Lee et al., 2022).

Economic disparities create barriers to accessing nutritious foods, influencing CMDs symptoms (Conte & De Luca, 2023). Nutritional patterns often correspond with a socioeconomic gradient. Low-income households rely on affordable, less nutrient-dense food options, like 100% fruit juice or white potatoes, which lack nutritional content and limit health benefits (Wallace et al., 2019). Similarly, leafy greens, consumed by higher-income households, offer nutrients like calcium and potassium, known to mitigate CMDs symptoms (Wallace et al., 2019). As a result, economic disparities directly impact dietary quality, thus influencing overall perceived health among adults with CMDs (Conte & De Luca, 2023).

Perceived health is another key indicator of overall well-being and health. Individuals experiencing extreme pain are 13 times more likely to report poor perceived health compared to those with little pain, who tend to report fair perceived health (Axon & Smith, 2024). In geriatric rehabilitation, results show older patients with a high risk for malnutrition were more likely to have poor perceived health than those at a low or medium risk (Söderhamn et al., 2007). Collectively, these findings suggest improper nutrition predicts the initial onset of CMD symptoms and perpetuates pain severity, contributing to poorer perceived health rates among individuals with CMDs.

While prior studies suggest diet plays a role in musculoskeletal health, pain severity, and perceived health, the complex relationship among these variables with household income remains unexplored (Putrik et al., 2015). Understanding this relationship is crucial for health promotion and developing targeted interventions around improving quality of life and health equity for individuals with CMDs (Goyal & Agrawal, 2021; Reginster, 2002). Bridging this gap can inform policies and community strategies, enabling underserved populations to adopt healthier dietary choices (Reginster, 2002). This alleviates suffering, promotes well-being, and lessens the long-term impact of CMDs (Reginster, 2002). To address these gaps, our study will explore the following research questions: 1. In Canadian adults over 35 years of age with a self reported chronic musculoskeletal disorder diagnosis (arthritis, osteoporosis, and fibromyalgia), is there an association between fruit and vegetable intake (low: <5 servings per day, moderate: 5-10 servings per day) and overall self reported perceived health scores (measured by poor, fair, good, very good, excellent)? 2. Among Canadian adults over the age of 35 with chronic musculoskeletal disorders (arthritis, osteoporosis, and fibromyalgia), is higher annual household income of $60,000 + associated with a higher fruit and vegetable intake (5-10 servings/day), while lower annual household income less than $59,999 is associated with lower intake (under 5 servings/day)?

Methods

Study Design and Database

This observational cross-sectional study took data from the Canadian Community Health Survey (CCHS), a nationwide cross-sectional survey conducted by Statistics Canada (Odesi, 2025). This study used data from the 2019 collection cycle from January to December

(Capaldi et al., 2021). The CCHS employs a multi-stage, stratified, clustered random sampling design ensuring demographic representativeness (Odesi, 2024). Trained personnel collected data via online and in-person interviews (Statistics Canada, 2024). Respondents could complete the interviews in English or French to help reduce language barriers (Statistics Canada, 2019). The database used a 24-hour recall to ensure response accuracy (Ahmed et al., 2021). Weighting protocols, including non-response adjustments, matched yearly national population estimates (CCHS, 2024). Sample weights like WTS_M ensured every response was weighted accordingly (CCHS, 2024). The database also used positive Poisson bootstrap weights for each duplicate, which were generated independently from one another (CCHS, 2024).

Study Participants

The analytical sample was derived from 10,000 respondents (Odesi, 2025). We included participants if they were above the age of 35 with a self-reported diagnosis of a CMDs of arthritis, osteoporosis, or fibromyalgia. We excluded individuals under the age of 35, those with incomplete or missing data on fruit and vegetable intake, household income, perceived health, and individuals with other conditions likely to influence perceived health. After applying inclusion/exclusion criteria, the final sample included 9,009 participants.

Study Outcomes

The primary outcome was self-reported perceived health, measured using a 5-point Likert scale (poor, fair, good, very good, and excellent). This scale is commonly used in population health research and has been linked to CMD morbidity risk (Lewis et al., 2011). Whitmore et al. (2020) noted this as the standard self-report measure among older adults, as it offers a patient-centered metric, emphasizing its relevance to our study.

The primary independent variable was daily fruit and vegetable intake, categorized as low (<5 servings/day) and high (5–10 servings/day). Intake was assessed using self-reported 24-hour recall, a method considered one of the gold standards of dietary intake research despite potential recall bias (Tanweer et al., 2022). Lafreniere et al. (2019) validated the web-based 24-hour dietary recall tool, confirming its reliability in evaluating vegetable and fruit intake among older adults. The secondary independent variable was total household income, categorized as below $59,999 a year and over $60,000 a year. This threshold aligns with Statistics Canada's income data from 2019-2020 (Statistics Canada, 2019). The median after-tax income of Canadian households during 2019-2020 was around $62,000, justifying this threshold for analysis (Statistics Canada, 2019).

The secondary outcome was dietary fruit and vegetable intake, analyzed by the same yearly household income brackets. Higher-income individuals consume healthier diets due to greater food access and nutrition knowledge (Riediger et al., 2007). Riediger et al. (2007) highlight roughly 60% of Canadian adolescents consume fruits and vegetables fewer than five times per day, with household yearly income influencing these dietary patterns.

Additional Data Collection

We included several demographic and contextual covariates to control for confounding variables. However, complete sociodemographic data was unavailable in the dataset, so we used several proxy variables instead. We included sex at birth (male/female) to examine sex-based differences in dietary intake and perceived health. We categorized age into three groups (35–49, 50–64, ≥65) to assess age-related variations. We included life satisfaction using a 5-point Likert scale from "very satisfied" to "very dissatisfied" as a substitute for psychological well-being, influencing perceived health. We included province of residence (limited to Prince Edward Island, Nova Scotia, Saskatchewan, and Alberta) to consider regional variation in dietary intake for CMD patients. These provinces are not representative of the entire population but of those selecting a CMD diagnosis on the survey. We included indigenous identity (First Nations, Métis, Inuk) for health outcomes and dietary access disparities. We used immigration status (Canadian-born, non-permanent resident) to determine the influence of differing nutritional habits on perceived health. No additional ethnicity or citizenship data were provided.

Statistical Analysis

We observed two types of missing data: complete and partial non-response from the respondents. Missing data values were reported as valid skip, don't know, refusal, or not stated and were deleted from the dataset and excluded from the study.

We performed data analyses using IBM SPSS Statistics Version 30. Descriptive statistics summarized baseline participants' characteristics (see additional data collection). We set statistical significance at $p < 0.05$ for all analyses, presenting categorical variables as frequencies and percentages. We used chi-square tests to test group differences in fruit and vegetable intake by household income and perceived health. We measured effect sizes for chi-square analysis using PHI coefficients. We drafted bar graph illustrations of associations using GraphPad Prism. We then conducted binary and multinomial logistic regression analyses to assess associations between fruit and vegetable intake, income, and perceived health, adjusting for demographic covariates. We assessed multicollinearity before model construction and included survey weights provided by Statistics Canada (see study design) to ensure adequate representative results.

Results

Study Sample

Our study initially included 10,000 cases. We excluded 991 participants due to incomplete data on fruit and vegetable intake, perceived health, or household income. Thus, our final study sample consisted of 9,009 study participants.

Statistical Analysis

We computed descriptive statistics to summarize participants' characteristics grouped by daily fruit and vegetable consumption: less than 5 servings (n=6,825) and between 5 and 10 servings (n=2,184) (Table 1). All data is categorical and is not in normal distribution.

Participants consuming between 5 and 10 fruit and vegetables per day were more likely to be female (74.8% vs. 62.4%, $p < 0.001$), report higher household incomes (≥ $80,000, 35.9% vs. 33.9%, $p = 0.004$), and be non-permanent residents (10.6% vs. 8.3%, $p = 0.001$). Moreover, consumption between 5 and 10 fruits and vegetables daily was also associated with better perceived health, with a greater proportion reporting "excellent" (11.3% vs. 7.8%, $p < 0.001$) or "very good" (34.2% vs. 31.6%, $p < 0.001$) and higher overall life satisfaction reporting "very satisfied (42.1% vs. 35.3%, $p < 0.001$).

Table 1

Baseline of study participant characteristics (N=9009).

Variables	Less than 5 fruit and vegetables/ day (n=6825) N (%)	Between 5 and 10 fruit and vegetables/day (n=2184) N(%)	P-value
Sex at birth			<0.001
Male	2569 (37.6)	551(25.2)	
Female	4256 (62.4)	1633(74.8)	
Age (years old)			0.85
35 to 49	190(2.8)	64(2.9)	
50 to 64	1061(15.5)	291(13.3)	
65 and older	522(80.9)	1814(83.1)	
Total household income			0.004
No income or less than $20,00	393(5.8)	83(3.8)	
$20,00 to $39,000	1729(25.3)	547(25.0)	
$40,000 to $59,000	1361(19.9)	422(19.3)	
$60,000 to $79,000	1026(15.0)	347(15.9)	
$80,000 or more	2316(33.9)	785(35.9)	
Perceived health			<0.001

Excellent	537(7.8)	264(11.3)	
Very good	2180(31.6)	748(34.2)	
Good	2357(34.1)	706(32.3)	
Fair	1310(19.0)	356(16.3)	
Poor	510(7.5)	128(5.9)	
Satisfaction with life in general*			<0.001
Very satisfied	2387(35.3)	913(42.1)	
Satisfied	3460(51.1)	1050(48.5)	
Neither satisfied or dissatisfied	594(8.8)	128(5.8)	
Dissatisfied	263(3.9)	59(2.7)	
Very Dissatisfied	65(1.0)	19(0.9)	
Province of residence of participant			<0.001
Prince Edward Island	1254(18.4)	329(15.1)	
Nova Scotia	1675(24.5)	460(21.1)	
Saskatchewan	1468(21.5)	567(26.0)	
Alberta	2428(35.6)	826(37.9)	
Aboriginal identity (First Nations/ Métis/ Inuk(Inuit))**			0.919
Yes	147(2.2)	46(2.2)	
No	6469(97.8)	2060(97.8)	
Immigrant***			0.001
Non-permanent resident	556(8.3)	225(10.6)	
Canadian born	6138(91.7)	1904(89.4)	

Note. Categorical data are expressed as n (%). P-values are derived from Chi-square tests.*less than 5 n=6769, between 5 and 10 n=2167, ** less than 5 fruit and vegetables/day n=6616, between 5 and 10 fruit and vegetables/day n=2106, ***less than 5 fruit and vegetables/day n=6694, between 5 and 10 fruit and vegetables/day n=2129

We performed Chi-square tests to examine the association between fruit and vegetable intake and self-reported perceived health. To further explore the relationship, we considered an ordinal logistic regression. However, the assumption of proportional odds was violated when we tested parallel lines ($p < 0.002$). As a result, we conducted a multinomial logistic regression instead. We assessed multicollinearity among predictor variables using Variance Inflation Factor (VIF) and tolerance values to confirm no multicollinearity was present. The covariates we included in the model were age, sex at birth, household income, immigration status, province of residence, and overall life satisfaction. We assessed model fit for the multinomial regression using the model chi-square statistic, -2 Log Likelihood, and Nagelkerke R^2.

To examine the relationship between household income and fruit and vegetable intake we first performed a Chi-square test. To further explore the relationship, we performed a linear logistic regression to assess multicollinearity among covariates using VIF and tolerance values before performing a binary logistic regression. We then conducted binary logistic regression to assess the likelihood of consuming 5-10 servings of fruits and vegetables per day vs. 5 servings, with household income as the main predictor. We controlled for the same covariates in this model as previously mentioned. We assessed model fit for the binary logistic regression using the Omnibus Test of Model Coefficients, -2 Log Likelihood, Hosmer-Lemeshow goodness of fit test, and Nagelkerke R^2.

A chi-square test revealed a significant association between fruit and vegetable intake and self-reported perceived health, with participants who consumed 5-10 servings of fruits and vegetables per day were more likely to rate their health as “very good” or “excellent” ($\chi^2 = 40.654$, df = 4, $\Phi = 0.067$, $p < 0.001$, N = 9009; Figure 1).

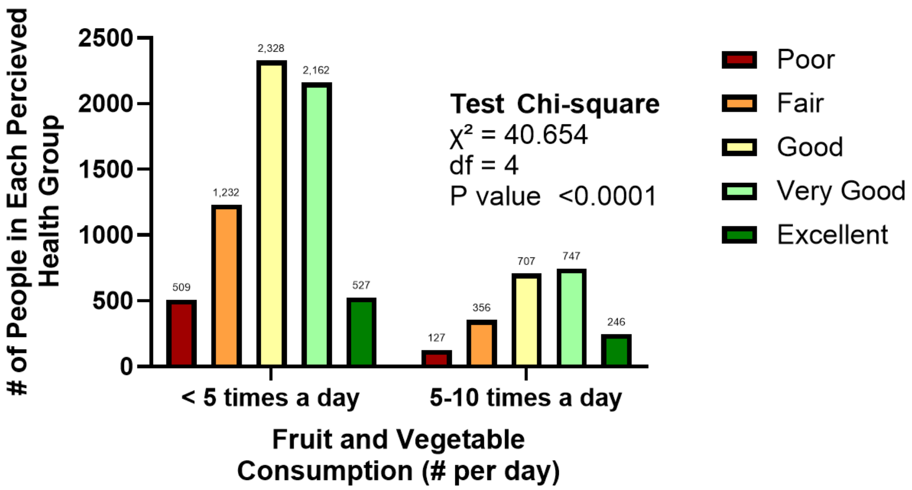


Figure 1. Group Comparison of Daily Fruit and Vegetable Consumption and Perceived Health Score. Relationship between daily fruit and vegetable consumption and perceived health status, comparing the number of individuals consuming daily fruit and vegetables <5 times per day with Poor (N=509), Fair (N=1232), Good (N=2328), Very Good (N=2162), Excellent (N=527) categories times per day or 5–10 times per day (N=127; 356; 707;747; 246). p-value: <0.001, suggesting a statistically significant positive association.

The overall multinomial logistic regression was statistically significant (χ^2 = 2139.76, df = 28, p < 0.001). The test had a modest explanatory power of Nagelkerke R^2 = 0.230. The multinomial logistic regression shows that when comparing to individuals with "poor" perceived health, those who consumed 5-10 servings of fruits and vegetables per day had significantly higher odds of reporting "excellent" health (B = -0.323 , SE = 0.142 , p < 0.001), "very good" (B = -0.105 , SE = 0.123, p < 0.001), and "good" health (B = 0.226 , SE = 0.037, p < 0.001). There was no significant difference in the odds of reporting "fair" health (B = -0.030, SE = 0.125 , p < 0.808).

A chi-square test indicated a significant but weak association between household income and fruit and vegetable intake, with participants with higher income (≥ $60,000) were more likely to consume 5-10 servings/day compared to those with lower income (< $60,000) ($\chi^2$ = 5.43, df = 1, Φ = 0.025 p = 0.020, N = 9009; figure 2).

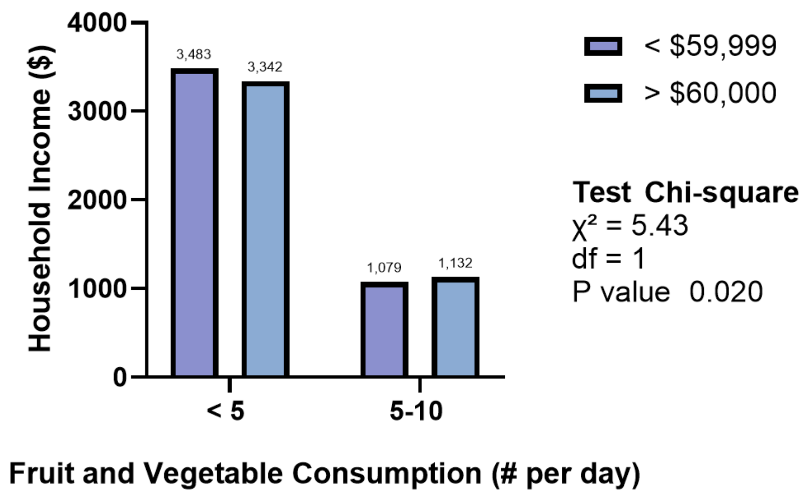


Figure 2. Group Comparison of Daily Fruit and Vegetable Consumption and Household Income. Relationship between daily fruit and vegetable consumption and household income, comparing the number of individuals in household income categories of <$ 59,999 daily fruit and vegetable intake of <5 times per day (N = 3483) vs. 5–10 times per day (N=1079) and >$ 60,000 (N= 3,342; 1,132). P-value: 0.0020, suggesting a statistically weak significant association.

The binary logistic model we used was significant (χ^2 = 229.45, df = 1, p < 0.001) and showed good fit using the Hosmer and Lemeshow test (χ^2 = 5.05, df = 8, p = 0.752). The test had a Nagelkerke R^2 = 0.039. The binary logistic regression test demonstrated that higher income was a statistically significant predictor of moderate (5-10 servings/day) fruit and vegetable intake (B = 0.115, SE = 0.053, p = 0.029).

Discussion

Understanding the elements that influence dietary intake and perceived health is critical for improving health outcomes for individuals with CMDs, as these conditions impact quality of life and overall physical functioning (Elma et al., 2020). This study examined the association between fruit and vegetable intake, household income, and perceived health among Canadian adults ages 35 years and older diagnosed with CMDs. We analyzed self-reported responses from 9,009 participants from the Canadian Health Measures Survey (CHMS), evaluating their perceived health, fruit and vegetable intake, and socioeconomic status scores. Our findings indicate individuals who consumed a moderate amount of fruit and vegetables (5-10 servings per day) were more likely to record greater perceived health than those who reported lower intake. Moreover, higher household income was positively associated with increased fruit and vegetable intake. However, the effect size was small.

Main Findings

Our results demonstrate a statistically significant association between fruit and vegetable intake and overall perceived health. Participants who consumed 5 to 10 servings per day were more likely to report their health as “very good” or “excellent” compared to those who consumed less than 5 servings per day, as shown in Figure 1. After accounting for socio-demographic factors through the multinomial logistic regression analysis, this association remains. Furthermore, participants consuming moderate amounts of fruit and vegetables had higher odds of reporting “excellent”, “very good”, or “good” health compared to those reporting “poor” overall health (Figure 1). While there was no significant difference between participants reporting “fair”, this trend reinforces the role of dietary habits within framing perceived health. However, the modest Nagelkerke R^2 value (0.230) shows other variables also significantly contribute to overall well-being.

Our analysis also revealed a statistically significant but weak association between household income and fruit and vegetable intake. Participants who reported a household income of $60,000 and above per year were more likely to consume at least 5 servings of fruit and vegetables per day than those with an income lower than $60,000, as shown in Figure 2. Yet, the effect size was small (Φ=0.025), and the Nagelkerke R^2 value was low (0.039), indicating that household income alone is not a strong predictor of fruit and vegetable intake. This shows other variables, such as food availability and cultural dietary habits, that may influence participants' fruit and vegetable intake.

Comparison with previous research

Consistent with previous research, our findings highlight that a higher daily intake of fruit and vegetables is associated with better overall perceived health. A literature review showed an increase of fruits within a diet reduces osteoarthritic knee pain and inflammation, indicating antioxidant-rich fruits may also provide relief for individuals experiencing CMDs-related pain. However, this review relied on numerous papers, which could introduce publication bias (Kurapatti & Carreira, 2023). Perna et al. (2020) found 96% of their participants with musculoskeletal pain had vitamin D deficiencies, while a study by Lewis et al (2019) found adults over the age of 65 had reduced muscle loss when consuming a diet with high potassium. These results emphasize the role of nutrition in maintaining musculoskeletal health and decreasing pain associated with CMDs. These modifiable factors impact an individual's perceived health and correlate with our study's findings.

A study from Darmon and Drewnowski (2008), found socioeconomic status also contributes to dietary habits, as higher-income individuals are more likely to consume nutrient-dense foods, however, their results may simplify the complex factors determining dietary choices. Although our study found a weak association between household income and fruit and vegetable intake, it points out other factors like food literacy, education, and food accessibility have a role in shaping individuals' dietary habits (Beaulac et al., 2009; Darmon & Drewnowski, 2008; Drewnowski, 2012). This highlights the need for policies that address affordability and accessibility within lower-income communities, focusing on individuals with CMDs.

Perceived health is influenced not only by pain severity within individuals with CMDs and their diet but also by broader social determinants. Poor nutrition has been linked with increased musculoskeletal pain, negatively impacting individuals' self-reported health (Bárbara Pereira Costa et al., 2016). However, this study only examined participants from Brazil, limiting the results' generalizability to broader demographics. Our study supports this, as results show individuals consuming 5 to 10 servings of fruit and vegetables daily had significantly higher odds of reporting “excellent” or “very good” health. These factors, however, were only shown to account for some variability within perceived health, suggesting other factors such as physical health, access to food, and psychological factors play a role in shaping how individuals perceive their overall health.

Limitations

While our findings support the research hypothesis, several limitations must be acknowledged. First, the reliance on self-report measures for fruit and vegetable intake and perceived health may introduce bias. Participants may overestimate or underestimate their fruit and vegetable intake due to either error in their recall or social desirability, notably because our data collection used standardized questionnaires either in-person or over the telephone, with the use of a 24-hour food recall, this design is subject to self-report bias and recall bias. Secondly, our study categorized fruit and vegetable intake into only two groupings, which may have concealed relationships between varying levels of intake and overall perceived health. Additionally, our analysis did not differentiate between types of fruit and vegetables, despite having varying nutritional benefits. While we accounted for socio-demographic factors, our measure of income may not accurately reflect participants' access to nutrient-dense foods, as exact locations were unavailable for analysis. Finally, our study did not consider other potential confounding variables such as physical activity, medication use, and additional health concerns.

Future Direction and Implications

The results of this study showcase the importance of health initiatives that promote fruit and vegetable intake within individuals with CMDs, particularly within lower-income populations where other dietary barriers exist. Nutrient-dense foods can improve perceived health outcomes, and strategies such as subsidizing fresh produce, expanding food markets and community gardens, and implementing educational programs about nutrition could support healthier choices at a more accessible and affordable cost.

This study addresses the critical gap in research on how fruit and vegetable intake impacts perceived health in individuals with CMDs and how household income affects dietary intake. Future research should explore the long-term impact of fruit and vegetable intake on overall perceived health through a longitudinal design. Utilizing objective measures to gather dietary information, such as food diaries, may reduce bias and improve data accuracy. Also, it would be valuable to examine other factors that may contribute to perceived health and fruit and vegetable intake, which this study did not capture. This can include physical activity through activity logs, medication use, and psychological factors through standardized questionnaires. Investigating potential barriers like food accessibility can further advise public health strategies to reduce disparities and promote well-being in such populations.

References

Ahmed, M., Praneet, Ng, A., & L'Abbe, M. R. (2021). Nutrient intakes of Canadian adults: Results from the Canadian Community Health Survey (CCHS)–2015 Public Use Microdata File. *The American Journal of Clinical Nutrition,* 114(3). https://doi.org/10.1093/ajcn/nqab143

Axon, D. R., & Smith, J. (2024). Relationship between pain severity and self-perceived health among united states adults: A cross-sectional, retrospective database study. *Medicine*, *103*(50). https://doi.org/10.1097/md.0000000000040949

Beaulac, J., Kristjansson, E., & Cummins, S. (2009). A systematic review of food deserts, 1966-2007. *Preventing chronic disease*, *6*(3), A105.https://pmc.ncbi.nlm.nih.gov/articles/PMC2722409/

Bárbara Pereira Costa, A., Andrade Carneiro Machado, L., Marcos Domingues Dias, J., Keller Coelho de Oliveira, A., Ude Viana, J., da Silva, S. L., Gonçalves Pereira Couto, F., Lustosa Torres, J., Mendes, L. P., & Correa Dias, R. (2016). Nutritional Risk is Associated with Chronic Musculoskeletal Pain in Community-dwelling Older Persons: The PAINEL Study. *Journal of nutrition in gerontology and geriatrics*, *35*(1), 43–51. https://doi.org/10.1080/21551197.2015.1125325

Canadian Community Health Survey (CCHS) -Annual Component. (2024). https://communitydata.ca/sites/default/files/2024-05/Apr30_2024_CDP_PUMF_Webinar.pdf

Capaldi, C. A., Liu, L., & Dopko, R. L. (2021). Positive mental health and perceived change in mental health among adults in Canada during the second wave of the COVID-19 pandemic. *Health Promotion and Chronic Disease Prevention in Canada*, 41(11). https://doi.org/10.24095/hpcdp.41.11.05

Conte, C., & De Luca, P. (2023). Editorial: Impact of diet-related disorders on musculoskeletal health. *Frontiers in Nutrition, 10.* https://doi.org/10.3389/fnut.2023.1301837

Darmon, N., & Drewnowski, A. (2008). Does social class predict diet quality? *The American Journal of Clinical Nutrition. 87*(5), 1107-1117. https://doi.org/10.1093/ajcn/87.5.1107

Drewnowski, A. (2012). The economics of food choice behavior: why poverty and obesity are linked. *Nestle Nutrition Institute workshop series*, *73*, 95–112. https://doi.org/10.1159/000341303

Elma, Ö., Yilmaz, S. T., Deliens, T., Clarys, P., Nijs, J., Coppieters, I., Polli, A., & Malfliet, A. (2020). Chronic musculoskeletal pain and nutrition: Where are we and where are we heading? *PM&R*, *12*(12), 1268–1278. https://doi.org/10.1002/pmrj.12346

Goyal, V., & Agrawal, M. (2021). Effect of supplementation of vitamin D and calcium on patients suffering from chronic non-specific musculoskeletal pain. *Journal of Family Medicine and Primary Care, 10*(5), 1839-1844. https://doi.org/10.4103/jfmpc.jfmpc_1699_20

Krsnich-Shriwise, S. (1997). Fibromyalgia syndrome: An overview. *Physical Therapy*, *77*(1), 68–75. https://doi.org/10.1093/ptj/77.1.68

Kurapatti, M., & Carreira, D. (2023). Diet composition’s effect on chronic musculoskeletal pain: A narrative review. *Pain Physician*, 26(6), 527-534. https://www.painphysicianjournal.com/current/pdf?article=Nzc1Nw%3D%3D&journal=156

Lafrenière, J., Couillard, C., Lamarche, B., Laramée, C., Vohl, M. C., & Lemieux, S. (2019). Associations between self-reported vegetable and fruit intake assessed with a new web-based 24-h dietary recall and serum carotenoids in free-living adults: a relative validation study. *Journal of Nutritional Science,* 8. https://doi.org/10.1017/jns.2019.23

Lee, S. H., Moore, L. V., Park, S., Harris, D. M., & Blanck, H. M. (2022). Adults meeting fruit and vegetable intake recommendations — United States, 2019. *MMWR. Morbidity and Mortality Weekly Report*, *71*(1), 1-9. https://doi.org/10.15585/mmwr.mm7101a1

Lewis, C., Anju Relan, O. Joe Hines, Tillou, A., & Hiatt, J. R. (2011). Morbidity and Mortality as a Televideoconference: A Randomized Prospective Evaluation of Learning and Perceptions. 212(3), 400–405. https://doi.org/10.1016/j.jamcollsurg.2010.12.002

Lewis, R., Gómez Álvarez, C. B., Rayman, M., Lanham-New, S., Woolf, A., & Mobasheri, A. (2019). Strategies for optimising Musculoskeletal Health in the 21st Century. *BMC Musculoskeletal Disorders*, *20*(1). https://doi.org/10.1186/s12891-019-2510-7

Nicklett, E. J., & Kadell, A. R. (2013). Fruit and vegetable intake among older adults: A scoping review. *Maturitas*, *75*(4), 305–312. https://doi.org/10.1016/j.maturitas.2013.05.005

Nisa, H., & Kurotani, K. (2023). Editorial: Diet Quality, socioeconomic differences, and health disparities. *Frontiers in Nutrition*, 10. https://doi.org/10.3389/fnut.2023.1250439

Odesi. (2025). *Canadian Community Health Survey, 2019-2020 Annual Component* https://odesi.ca/en/details?id=/odesi/doi --- 10-5683_SP3_ZVCGBK.xml

Perna, S., Alalwan, T. A., Al-Thawadi, S., Negro, M., Parimbelli, M., Cerullo, G., Gasparri, C., Guerriero, F., Infantino, V., Diana, M., D'Antona, G., & Rondanelli, M. (2020). Evidence-based role of nutrients and antioxidants for chronic pain management in musculoskeletal frailty and sarcopenia in aging. *Geriatrics*, *5*(1), 16. https://doi.org/10.3390/geriatrics5010016

Putrik, P., Ramiro, S., Chorus, A. M., Keszei, A. P., & Boonen, A. (2015). Socioeconomic inequities in perceived health among patients with musculoskeletal disorders compared with other chronic disorders: Results from a cross-sectional Dutch study. *RMD Open, 1*(1). https://doi.org/10.1136/rmdopen-2014-000045

Rachner, T. D., Khosla, S., & Hofbauer, L. C. (2011). Osteoporosis: Now and the future. *The Lancet*, *377*(9773), 1276–1287. https://doi.org/10.1016/s0140-6736(10)62349-5

Reginster, J. -Y. (2002). The prevalence and burden of arthritis. *Rheumatology*, *41*(suppl_1), 3–6. https://doi.org/10.1093/rheumatology/41.s1.3

Riediger, N. D., Shooshtari, S., & Moghadasian, M. H. (2007). The Influence of Sociodemographic Factors on Patterns of Fruit and Vegetable Consumption in Canadian Adolescents. *Journal of the American Dietetic Association*, 107(9), 1511–1518. https://doi.org/10.1016/j.jada.2007.06.015

Statistics Canada. (2021, March 23). *The Daily — Canadian Income Survey, 2019*. https://www150.statcan.gc.ca/n1/daily-quotidien/210323/dq210323a-eng.htm

Söderhamn, U., Bachrach-Lindström, M., & Ek, A. (2007). Nutritional screening and perceived health in a group of Geriatric Rehabilitation Patients. *Journal of Clinical Nursing*, *16*(11), 1997–2006. https://doi.org/10.1111/j.1365-2702.2006.01805.x

Tanweer, A., Khan, S., Mustafa, F. N., Imran, S., Humayun, A., & Hussain, Z. (2022). Improving dietary data collection tools for better nutritional assessment. *Computer Methods and Programs in Biomedicine Update*, 2, 100067. https://doi.org/10.1016/j.cmpbup.2022.100067

Wagenaar, C. A., van de Put, M., Bisschops, M., Walrabenstein, W., de Jonge, C. S., Herrema, H., & van Schaardenburg, D. (2021). The effect of dietary interventions on chronic inflammatory diseases in relation to the microbiome: A systematic review. *Nutrients*, *13*(9), 3208. https://doi.org/10.3390/nu13093208

Wallace, T. C., Bailey, R. L., Blumberg, J. B., Burton-Freeman, B., Chen, C. O., Crowe-White, K. M., Drewnowski, A., Hooshmand, S., Johnson, E., Lewis, R., Murray, R., Shapses, S. A., & Wang, D. D. (2019). Fruits, vegetables, and health: A comprehensive narrative, umbrella review of the science and recommendations for

Enhanced Public Policy to improve intake. *Critical Reviews in Food Science and Nutrition*, *60*(13), 2174–2211. https://doi.org/10.1080/10408398.2019.1632258

Whitmore, C., Markle-Reid, M., Fisher, K., McAiney, C., & Ploeg, J. (2020). What Factors Shape Self-Reported Health Among Community-Dwelling Older Adults? A Scoping Review. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, 40(2), 177–192. https://doi.org/10.1017/s0714980820000124

Exploring the Association Between Different Modes of Delivery and Maternal Age, with Neonatal Sepsis among Young Infants in Dhaka, Bangladesh: A Cross-Sectional Analysis

Alessandra Giglia and Vicki Yiannakis

McMaster University

School of Interdisciplinary Science

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8, 2025

Exploring the Association Between Different Modes of Delivery and Maternal Age, with Neonatal Sepsis among Young Infants in Dhaka, Bangladesh: A Cross-Sectional Analysis

Introduction

Neonatal sepsis is an illness caused by severe or life-threatening infection in young infants, driven by an immature immune system (Molloy et al., 2020). It remains an urgent global health concern, particularly in low and middle-income countries (LMICs), where the burden of neonatal sepsis is disproportionately high (Milton et al., 2022; Perin et al., 2022). According to Fleischmann et al. (2018), infants have the highest incidence of sepsis among age groups, with 3 million cases globally each year and a mortality rate of 19% (Molloy et al., 2020). Neonatal sepsis places a significant strain on the global healthcare system, amounting to a cost of $670 billion CAD annually (Ranjeva et al., 2018).

Neonatal sepsis can be classified as early-onset sepsis (EOS) or late-onset sepsis (LOS) (Shane et al., 2017). EOS results from mother-to-infant transmission and occurs within the first 72 hours of life (Rubio-Mora et al., 2025). LOS is transmitted to the infant from environmental pathogen exposure, and occurs after the EOS time period, within the first 28 days of life (Hayes et al., 2023). The infant adaptive immune system is tolerogenic and anti-inflammatory, and infants rely heavily upon innate immunity (Raymond et al., 2017). As a result, their immature immune system is especially vulnerable to pathogens (Singh et al., 2025). The inadequate local response to an infection can then provoke a dysregulated, systemic, proinflammatory cascade that leads to neonatal sepsis (Raturi & Chandran, 2024).

Neonatal sepsis has been linked to a range of adverse outcomes for infants. A meta-analysis by Ong et al. (2024) found infants with a history of sepsis are at an increased risk for cognitive delays, hearing and visual impairments, and delayed psychomotor development. Given these serious consequences, understanding the risk factors for neonatal sepsis is critical. Among these, maternal factors like mode of delivery and maternal age have emerged as significant risk factors. A meta-analysis in Saudi Arabia found infants born via cesarean section (CS) were 2.13 times more at risk for neonatal sepsis. (Hamid et al., 2021). This may be because vaginal birth exposes neonates to maternal microbiota, aiding microbiome establishment and immune development (Björkstén, 2004; Malamitsi-Puchner et al., 2005). While cesarean delivery limits this exposure, it often leads to colonization by hospital-associated bacteria and delayed breastfeeding, which may reduce access to colostrum's immunological benefits (Dominguez-Bello et al., 2010). In contrast, a hospital-based observational study in Indore, India reported that vaginal delivery was associated with a 2.29 times higher risk of sepsis compared to CS, though this finding may reflect specific hospital conditions (Nath et al., 2016).

Similarly, maternal age has been associated with neonatal sepsis risk. A retrospective cohort study by Wu et al. (2024) concluded that neonatal sepsis risk increases by 2.16 times for mothers younger than 26 years old. This increased risk may be due to younger mothers having less education, lower economic status, and less knowledge about potential warning signs during pregnancy (Wu et al., 2024). Contrastingly, an ICU-based case control study in Egypt concluded that infants born to mothers older than 35 years old were 1.6 times more likely to have sepsis (Salama & Tharwat, 2023). Older mothers face greater predisposition to gestational hypertension, diabetes, congenital abnormalities, and premature birth, all which can weaken infants’ immunity (Levenson et al., 2020).

Despite evidence linking maternal factors like delivery mode and maternal age to neonatal sepsis risk, findings are inconsistent across populations and healthcare settings. While some studies link CS to increased neonatal sepsis risk, others report a higher risk after vaginal delivery. Similarly, both younger and advanced maternal ages have been linked to increased neonatal sepsis risk (Salama & Tharwat, 2023; Wu et al., 2024). These inconsistencies create confusion in understanding how these maternal factors influence neonatal sepsis, particularly as healthcare settings and population characteristics vary widely. In LMICs, neonatal sepsis stands out as the leading cause of neonatal mortality, contributing to 98% of all neonatal deaths (Rahman et al., 2021). Given this, it is important to examine specific LMICs where the burden is most pronounced. In Bangladesh in particular, neonatal sepsis accounts for 9% of all newborn deaths, however there are currently no studies that examine the relationship between maternal risk factors and neonatal sepsis risk in this context. With Bangladesh CS rates increasing by 27% over the past decade and an average maternal age of 18 at first birth, studying these associations in this setting is crucial (Hossain et al., 2022; Sizear & Rashid, 2024)

This study aims to fill this gap by examining how mode of delivery and maternal age are associated with the incidence of neonatal sepsis in Dhaka, Bangladesh, using observational cohort study data from Synbiotics for Early Prevention of Severe Infections in Infants (SEPSiS) (Fung et al., 2024). The primary outcome, neonatal sepsis, is defined in the SEPSiS study as physician-documented clinical signs or a sepsis/serious infection diagnosis, plus a positive sterile culture or hospitalization with $\geq$ 5 days of antibiotics (Fung et al., 2024). Specifically, this study seeks to answer: Among infants aged 0-60 days in Dhaka, Bangladesh, during the period from November 25, 2020, to February 18, 2022: What is the association between delivery by cesarean section, compared to vaginal birth, and neonatal sepsis incidence? What is the correlation between maternal age and neonatal sepsis incidence?

Methods

Study Design and Database

This was a cross-sectional, observational study conducted on open-source data provided by the Synbiotics for the Early Prevention of Severe Infections in Infants (SEPSiS) observational cohort study (Fung et al., 2024). The original purpose of the SEPSiS database was to provide baseline data to inform future clinical trials aimed at assessing the effectiveness of synbiotics in preventing severe infections, in low-resource settings, and to examine the effects of variations in how sepsis is defined on incidence estimates. Between November 25, 2020, and February 18, 2022, a total of 1939 generally healthy infants born in Dhaka, Bangladesh, were enrolled in the study and monitored over a 60-day observational period. Sepsis cases were identified through home visits or caregiver self-referral.

Study Participants

Eligible participants were screened by study personnel for eligibility at two healthcare centres in Dhaka: the Maternal and Child Health Training Institute (MCHTI) and Mohammadpur Fertility Services and Training Centre (MFSTC). To be enrolled in the study, infants had to be between 0 days (date of birth) and 4 days of age, delivered at the study hospital, and have informed consent provided by a parent or guardian. Infants were excluded if they were not orally feeding, were undergoing mechanical ventilation or cardiac support, had a major congenital anomaly of the gastrointestinal tract, were receiving parenteral antibiotics at the time of enrollment, were born at a weight of less than 1500 g, or if death/major surgery was probable within the first week of life. Infants were required to be generally healthy upon enrollment to reliably assess neonatal sepsis risk factors without interference from pre-existing conditions that could bias the outcomes. Since the original purpose of this dataset was to inform future clinical trials on severe infection prevention using synbiotics, participants were also excluded if the mother was receiving any prenatal or postpartum non-dietary probiotic supplements, or if the infant received any postnatal administration of non-dietary probiotic or prebiotic supplements. Cases with missing data for mode of delivery, maternal age, and neonatal sepsis were excluded from our dataset via pairwise deletion.

Outcomes

The primary outcome of this study was neonatal sepsis incidence. Infants were classified as ‘yes’ or ‘no’ for sepsis occurrence within the 60-day timeframe. Neonatal sepsis was defined as at least one sign of clinical severe infection (poor feeding, severe chest indrawing, fever [≥38°C], hypothermia [<35.5°C], or lethargy) documented by a study medical officer or a non-study treating physician’s diagnosis of sepsis or a serious bacterial infection. Additionally, the case had to meet either at least one of the following two criteria: A non-study physician’s decision to admit the infant to the hospital, with at least one dose of a parenteral antibiotic given on the first day and intended or actual treatment for at least five consecutive days; and/or a positive blood or cerebrospinal fluid culture for a pathogenic bacterial or fungal organism. This case definition was designed to balance permissiveness, stringency, and feasibility of operation in the study setting. While a positive sterile site culture is regarded as the gold standard when diagnosing infant sepsis, sick infants in LMICs are often assessed in the community or at first-level health facilities where access to microbiological resources are limited (Yadav & Yadav, 2022). This case definition combines clinical signs, antibiotic treatment and microbiological criteria to create a precise and feasible definition. Infants were monitored for sepsis from enrollment to 60 days of age, as this is the period during which the World Health Organization defines a young infant (World Health Organization, 2015). Up to 12 clinical assessments were conducted by trained community health research workers (CHRWs) through in-person home or hospital visits, or by phone when necessary. During each assessment, CHRWs evaluated infants for signs of possible sepsis infection. Caregivers were encouraged to report concerning signs to the study team between visits. Infants showing signs of illness were referred to a study hospital for evaluation by a medical officer.

Additional Data Collection

Maternal age was determined during enrollment when a nurse assessing baseline characteristics asked the infant's mother, 'What is your age in years?' Mode of delivery was verified using the infant's medical record. Additional demographic and clinical characteristics, presented in **Table 1.0,** were collected for descriptive purposes to provide a better understanding of the study population. These baseline characteristics were compared based on mode of delivery (cesarean section vs. vaginal birth). Infant characteristics included infant sex, birth weight, feeding pattern, gestational age, and infant age at enrollment. Maternal characteristics included maternal education, parity, and administration of peripartum antibiotics.

Statistical Analyses

Continuous variables (birth weight, infant age, and maternal age) were assessed for normality using the Shapiro-Wilk test. For Table 1.0, continuous variables were summarized using median (IQR) and categorical variables were summarized using frequency (%). Baseline characteristics by mode of delivery were compared using a Chi-square test for categorical variables and a Mann-Whitney U-test for continuous variables, as they were not normally distributed. Since both the independent variable (mode of delivery) and the dependent variable (sepsis incidence) were categorical, the association was analyzed by comparing the proportion of infants with sepsis (yes vs. no) between vaginal and cesarean deliveries using a 2×2 Chi-square test. Yates' continuity correction was applied to adjust for small expected frequencies. The association between maternal age and sepsis incidence was evaluated using Spearman's rank correlation coefficient, as maternal age was not normally distributed. All analyses were performed using SPSS (Version 30.0), with statistical significance set at $p < 0.05$.

Results

Assessing normality table

Variable	**Shapiro-Wilk significance value**	**Normal or not normal?**	**Measure of central tendency to be reported**
BirthWeight	<0.001	Not Normal	Median and IQR
Infant Age at Enrollment	<0.001	Not Normal	Median and IQR
Maternal Age	<0.001	Not Normal	Median and IQR

Table 1.

Baseline Infant and Maternal Characteristics of Participants, Overall and By Mode of Birth.

Characteristic	**Overall**	**C-section**	**Vaginal Birth**	**p-Value**
INFANT CHARACTERISTICS				
Number of Infants, n (%)	1939 (100)	1064 (54.9)	875 (45.1)	
Gestational age at delivery, n (%)				0.541

Term (≥ 37 weeks)	1769 (91.2)	972 (92.6)	797 (91.7)	
Pre-term (<37 weeks)	150 (7.7)	78 (7.4)	72 (8.3)	
Infant Sex, n (%)				0.444
Male	926 (47.8)	517 (48.6)	409 (46.7)	
Female	1013 (52.2)	547 (51.4)	466 (53.3)	
Birth weight (grams)[a]	2860 (394)	2900 (526)	2805 (535)	**<0.001**
Feeding pattern at or near enrolment, n (%)				0.057
Exclusively breastfed	1843 (95.0)	1001 (94.5)	842 (96.4)	
Not exclusively breastfed or not breastfed	89 (4.6)	58 (5.5)	31 (3.6)	
Infant age at enrolment (days)[b]	1.00 (2.0)	2.00 (1.0)	0.00 (1.0)	**<0.001**
MATERNAL CHARACTERISTICS[b]				
Age (years)[a]	24 (7)	25 (7)	22 (6)	**<0.001**
Maternal education, n (%)				**0.040**
Secondary complete or higher	769 (39.7)	449 (42.2)	323 (36.6)	
Secondary Incomplete	621 (32.0)	327 (30.8)	294 (33.6)	
None up to complete primary school	548 (28.3)	287 (27.0)	261 (29.8)	

Parity[c]				**<0.001**
First live birth, n (%)	832 (42.9)	397 (37.3)	435 (49.7)	
Maternal peripartum antibiotics administered, n (%)				**<0.001**
None	93 (4.8)	42 (4.0)	51 (5.8)	
Intrapartum[d] only	97 (5.0)	93 (8.7)	4 (0.5)	
Postpartum only	1425 (73.5)	632 (59.5)	793 (90.6)	
Both intrapartum and postpartum	323 (16.7)	296 (27.8)	27 (3.1)	

Note. Variables with missing data were gestational age at delivery, feeding pattern at or near enrolment, maternal age, maternal education, parity, and maternal peripartum antibiotics administered. Bolded p-values indicate statistical significance ($p<0.05$).

[a] Reported as median (IQR).

[b] The number of mothers for the 1939 infants is n = 1926.

[c] Parity refers to the mother’s total number of pregnancies minus the number of previous abortions/miscarriages that occurred during the first or second trimester of pregnancy.

[d] Intrapartum refers to antibiotics that were administered during labour, but prior to delivery.

Population Characteristics

One mother-infant pair was excluded via pairwise deletion, due to missing data for maternal age. A total of 1,939 infant-mother pairs were included in this study, with 54.9% infants (n = 1,064) delivered via CS and 45.1% (n = 875) via vaginal birth. Baseline characteristics were compared between infants delivered by CS and vaginally **(Table 1.0).** There were no significant differences in gestational age or infant sex. However, infants born via CS had a significantly higher median birth weight (2900 g) compared to vaginal birth (2805 g). Mothers who had a CS were significantly older than those who delivered vaginally (25 vs. 22 respectively). First-time births were significantly more common in the vaginal birth group (49.7% vs. 37.3% respectively). Maternal education differed significantly, with 42.2% of cesarean mothers completing secondary education or higher compared to 36.6% in the vaginal birth group.

There was no significant difference in infant feeding patterns at enrollment, with the majority of infants in both groups being exclusively breastfed. However, peripartum antibiotic use varied significantly by delivery mode. Among vaginal deliveries, 90.6% received only postpartum antibiotics, whereas 27.8% of cesarean births received both intrapartum and postpartum antibiotics.

Primary analysis

Of the 875 infants delivered vaginally, 63 (7.2%) developed sepsis, compared to 65 out of 1064 (6.1%) infants delivered by CS, as shown in **Figure 1.0**. While there was a slightly higher proportion of sepsis cases in infants delivered via vaginal birth, the Chi-square test indicated no significant difference in the incidence of sepsis between the two groups ($p = 0.384$), suggesting no association between mode of delivery and sepsis in the SEPSiS cohort.

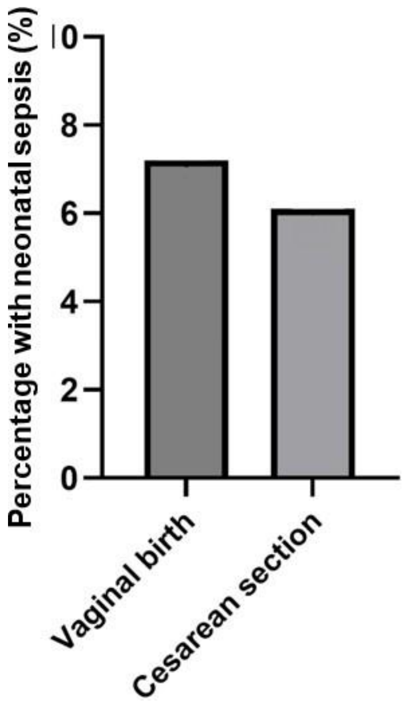


Figure 1.0. Proportion of neonatal sepsis cases across mode of delivery. Infants born via vaginal birth (7.2%) showed no significant increase in the proportion of infants with sepsis, compared to infants delivered via cesarean section (6.1%) (p = 0.384) .

Secondary analysis

Our secondary analysis of the SEPSiS cohort suggested no correlation between maternal age and incidence of neonatal sepsis in the infant (Spearman's $\rho = -0.021$, $p = 0.357$). For the 128 infants who had neonatal sepsis, the median maternal age was 23 years, and for the 1810 infants who had no sepsis, the median maternal age was 24 years. As depicted in **Figure 2.0**, the distribution of maternal age appears to be similar across both groups, indicating a lack of correlation between maternal age and sepsis incidence in this cohort.

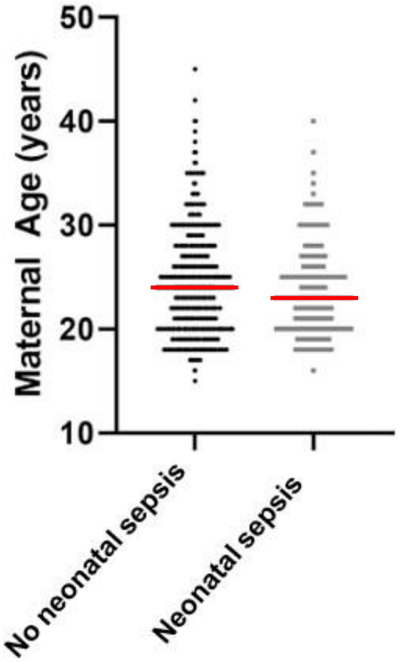


Figure 2. Distribution of maternal ages among infants with and without neonatal sepsis. Each dot represents one subject, and group medians are represented by a solid red line. The median for the non-sepsis group was 24 years of age, and the median for the sepsis group was 23 years of age. There was no significant difference ($p = 0.357$) in distribution of age in the sepsis group ($n = 128$) and non-sepsis group ($n = 1810$).

Discussion

Our study revealed no significant difference in neonatal sepsis incidence between cesarean and vaginal deliveries, with rates of 6.1% and 7.2%, respectively. Furthermore, we did not detect any significant correlation between maternal age and neonatal sepsis incidence. Previous studies have reported conflicting findings regarding delivery mode and maternal age as maternal risk factors for neonatal sepsis. For instance, a meta-analysis by Hamid et al. (2021) based in Saudi Arabia, reported infants born via CS were at 2.13x greater risk for neonatal sepsis. Meanwhile, Adatara et al. (2019) found mode of delivery was a significant predictor of neonatal sepsis, with CS accounting for 65% of neonatal sepsis cases. These studies support the proposition that cesarean delivery may predispose infants to sepsis due to altered microbiota colonization or delayed breastfeeding initiation (Dominguez-Bello et al., 2010; Malamitsi-Puchner et al., 2005). Whereas, Mehar et al., (2016) found that infants born via vaginal delivery in Indore, India, were 2.29x more likely to develop neonatal sepsis compared to those born via CS; this finding was attributed to specific conditions and practices unique to their hospital setting. Furthermore, previous literature suggests a correlation between infant sepsis and maternal age, although variability exists in the findings. A case control study by Salama and Tharwat (2023) reported a maternal age greater than 35 years old is linked to a 60% increased risk in neonatal sepsis. Whereas retrospective case control by Wu et al. (2024) found infants of mothers younger than 26 years of age were found to have a 2.16x greater risk of developing neonatal sepsis. While infants born to older mothers may face increased infection risk due to greater risk of pregnancy complications, younger mothers may be less educated, less knowledgeable about pregnancy-related warning signs, and have a lower socioeconomic status (Levenson et al., 2020; Wu et al., 2024).

Several factors may explain why we did not observe a significant association between either mode of delivery or maternal age and the incidence of neonatal sepsis. The exclusion of high-risk, low-birth-weight infants resulted in an overall healthier birth cohort, potentially lowering sepsis risk across all delivery modes (Belachew & Tewabe, 2020). Notably, CS-delivered infants in our study had significantly higher birth weights, which has been associated with lower susceptibility to neonatal infections (Belachew & Tewabe, 2020). In addition, there was a significant proportion of younger mothers; the median maternal age was 24 years old and 84% of mothers were under 30 years of age. The relative youth of the maternal population may have lessened age-related pregnancy complications and resulted in healthier births (Salama &

Tharwat, 2023). Peripartum antibiotic administration was also significantly more common among mothers who delivered via CS. Maternal antibiotics reduce bacterial load during delivery, thereby decreasing vertical transmission of pathogens to neonates (Chan et al., 2013). This is consistent with findings from Chan et al., (2014) who found neonatal sepsis incidence decreased from 1.7 to 0.6 per 1,000 live births following maternal intrapartum antibiotic use in Dhaka, Bangladesh.

Our study had several limitations to consider. The number of infants with sepsis was relatively small (66 per 1000 births), which may have limited our ability to detect statistically significant differences. In comparison, the BARNARDS cohort study reported a higher sepsis incidence at 166 per 1000 births across seven LMICs (Milton et al., 2022). This disparity may be attributed to differences in inclusion criteria: the BARNARDS study included EOS cases, while our study's enrollment window may have excluded infants who developed EOS within the first 72 hours of life. Additionally, infants in our study were recruited up until four days of age, and those receiving parenteral antibiotics were excluded. Therefore, our sepsis incidence estimate is not representative of a true birth cohort, which limits the generalizability of our findings to healthier infants. Additionally**,** the SEPSiS study spanned a 60-day observational period, during which infants primarily remained at home. Postnatal factors, such as hygiene practices, environmental exposures, and access to healthcare, may influence sepsis risk during this time, complicating efforts to associate sepsis incidence with maternal risk factors (Blencowe et al., 2011). It should also be noted that the SEPSiS cohort study occurred during the Covid-19 pandemic; pandemic lockdowns may have resulted in decreased health-seeking behaviour and identification of infections (Fung et al., 2024; Rizwan Khan et al., 2024). Whereas social distancing measures could have decreased infection rates (McBride, 2021).

This cross-sectional study had a number of strengths. For one, the sepsis case definition effectively balanced permissiveness and stringency. By integrating clinical criteria, physician diagnosis of sepsis, and at least one of the following – administration of antibiotics or blood culture confirmation – the study ensured a robust process for case identification. Additionally, the inclusion of two study centers (MCHTI and MFSTC) helped reduce the likelihood of site-specific biases, and allowed for better representation of different clinical practices in Dhaka, Bangladesh. Furthermore, continuous monitoring of infants throughout the study period ensured that cases were accurately detected and minimized the risk of missing cases.

Future studies should revise exclusion criteria to include low-birth-weight infants and infants receiving parenteral antibiotics, so that findings are more representative of a broader neonatal population. As well as, recruiting infants on day 0 of life would allow to improve the ability to capture EOS cases. Repeating a similar study post-Covid-19 pandemic may also enhance external validity by reflecting current health conditions and practices. Furthermore, there is a lack of studies that evaluate how maternal risk factors influence time of sepsis onset (EOS or LOS). In the future, a longitudinal study could effectively analyze this relationship to help provide direction for analyzing mechanisms behind EOS and LOS pathophysiology. Since more than half of Bangladeshi live in rural areas, where healthcare access and quality is limited, expanding this project to rural health centers could offer perspectives unique from urban environments (Akter & Kabir, 2023).

Overall, these findings suggest that neonatal sepsis risk is likely influenced by a complex interplay of factors beyond mode of delivery and maternal age alone.

References

Adatara, P., Afaya, A., Salia, S. M., Afaya, R. A., Konlan, K. D., Agyabeng-Fandoh, E.,…Boahene, I. G. (2019). Risk Factors Associated with Neonatal Sepsis: A Case Study at a Specialist Hospital in Ghana. *The Scientific World Journal*(1), 5. https://doi.org/10.1155/2019/9369051

Belachew, A., Tewabe, T., Belachew, A., & Tewabe, T. (2020). Neonatal sepsis and its association with birth weight and gestational age among admitted neonates in Ethiopia: systematic review and meta-analysis. *BMC Pediatrics 2020 20:1*, *20*(1). https://doi.org/10.1186/s12887-020-1949-x

Björkstén, B. (2004). Effects of intestinal microflora and the environment on the development of asthma and allergy. *Springer Nature*, *25*(3), 257–270. https://doi.org/10.1007/s00281-003-0142-2

Blencowe, H., Cousens, S., Mullany, L. C., Lee, A. C., Kerber, K., Wall, S.,…Lawn, J. E. (2011). Clean birth and postnatal care practices to reduce neonatal deaths from sepsis and tetanus: a systematic review and Delphi estimation of mortality effect. *BMC Public Health*, *11*(Suppl 3). https://doi.org/10.1186/1471-2458-11-S3-S11

Chan, G. J., Stuart, E. A., Zaman, M., Mahmud, A. A., Baqui, A. H., & Black, R. E. (2014). The effect of intrapartum antibiotics on early-onset neonatal sepsis in Dhaka, Bangladesh: a propensity score matched analysis. *BMC Pediatrics, 14*(1), 2-8. https://doi.org/10.1186/1471-2431-14-104

Chan, G., Lee, A., Baqui, A., J, J. T., & Black, T. (2013). Risk of early-onset neonatal infection with maternal infection or colonization: a global systematic review and meta-analysis - PubMed. *PLoS medicine, 10*(8). https://doi.org/10.1371/journal.pmed.1001502

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N.,…Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences*, *107*(26), 11971-11975. https://doi.org/10.1073/pnas.1002601107

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N.,…Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences*, *107*(26), 11971-11975. https://doi.org/10.1073/pnas.1002601107

Hamid, N. F., Ebrahim, M. E. E., Alsharari, N. M. T., Alyasi, A. S. M., & Rizk, M. M. M. (2021). Neonatal sepsis in association with birth weight, gestational age, and mode of delivery in Saudi Arabia: A systematic review and meta-analysis. *Medical Science, 25*(111), 1071-1079.

Hossain, M. A., Jahan, I., Haque, M. M., Sarwar, N., Yeasmin, N., & Islam, D. (2022). Rising Trends of Cesarean Section in Bangladesh: Associated Factors and Long-Term Complications on Health of Mother and Children. *Journal of Maternal and Child Health*, *7*(5), 532-542.

Kan, Bernard, R., Hamid Reza, & Lavoie, P. M. (2016). *An Immunological Perspective on Neonatal Sepsis: Trends in Molecular Medicine*. https://www.cell.com/trends/molecular-medicine/fulltext/S1471-4914(16)00031-9

Levenson, D., Romero, R., Garcia-Flores, V., Miller, D., Xu, Y., Sahi, A., ... & Gomez-Lopez, N. (2020). The effects of advanced maternal age on T-cell subsets at the maternal–fetal interface prior to term labor and in the offspring: a mouse study. *Clinical & Experimental Immunology*, *201*(1), 58-75.

Malamitsi-Puchner, A., Protonotariou, E., Boutsikou, T., Makrakis, E., Sarandakou, A., & Creatsas, G. (2005). The influence of the mode of delivery on circulating cytokine concentrations in the perinatal period. *Early Human Development*, *81*(4), 387-392. https://doi.org/10.1016/j.earlhumdev.2004.10.017

Mehar, V., Agarwal, S., Singh, R., Agarwal, A., Agrawal, N., & Majethia, A. (2016). Relationship between gestational age and mode of delivery with neonatal septicemia. *International Journal of Contemporary Pediatrics*, *3*(3), 891–895. https://doi.org/10.18203/2349-3291.ijcp20162361

Milton, R., Gillespie, D., Dyer, C., Taiyari, K., Carvalho, M. J., Thomson, K., Sands, K., Portal, E. A. R., Hood, K., Ferreira, A., Hender, T., Kirby, N., Mathias, J., Nieto, M., Watkins, W. J., Bekele, D., Abayneh, M., Solomon, S., Basu, S., … Chan, G. J. (2022). Neonatal sepsis and mortality in low-income and middle-income countries from a facility-based birth cohort: An international multisite prospective observational study. *The Lancet Global Health*, *10*(5), e661–e672. https://doi.org/10.1016/S2214-109X(22)00043-2

Molloy, E. J., Wynn, J. L., Bliss, J., Koenig, J. M., Keij, F. M., McGovern, M., Kuester, H., Turner, M. A., Giannoni, E., Mazela, J., Degtyareva, M., Strunk, T., Simons, S. H. P., Janota, J., Plotz, F. B., van den Hoogen, A., de Boode, W., Schlapbach, L. J., & Reiss, I. K. M. (2020). Neonatal sepsis: Need for consensus definition, collaboration and core outcomes. *Pediatric Research*, *88*(1), 2–4. https://doi.org/10.1038/s41390-020-0850-5

Nath, K. S., Ahmad, G. S., & Ali, M. H. (2016). An Observational Study of High Risk New Born: A Hospital Based Study. *Journal of Medical Science and Clinical Research*, *4*(7), 11527-11532.

Ong, W. J., Seng, J. J. B., Yap, B., He, G., Moochhala, N. A., Ng, C. L.,…Chong, S.-L. (2024). Impact of neonatal sepsis on neurocognitive outcomes: a systematic review and meta-analysis. *BMC Pediatrics*, *24*(1), 505. https://doi.org/10.1186/s12887-024-04977-8

Perin, J., Mulick, A., Yeung, D., Villavicencio, F., Lopez, G., Strong, K. L., Prieto-Merino, D., Cousens, S., Black, R. E., & Liu, L. (2022). Global, regional, and national causes of under-5 mortality in 2000–19: An updated systematic analysis with implications for the Sustainable Development Goals. *The Lancet Child & Adolescent Health, 6*(2), 106–115. https://doi.org/10.1016/S2352-4642(21)00311-4

Phelan, H., Yates, V., & Lillie, E. (2022). Challenges in healthcare delivery in low- and middle-income countries. *Anaesthesia and Intensive Care Medicine, 23*(8), 501–504. https://doi.org/10.1016/j.mpaic.2022.05.004

Rahman, A. E., Hossain, A. T., Siddique, A. B., Jabeen, S., Chisti, M. J., Dockrell, D. H., ... & El Arifeen, S. (2021). Child mortality in Bangladesh–why, when, where and how? A national survey-based analysis. *Journal of global health*, *11*, 04052.

Ranjeva, S. L., Warf, B. C., & Schiff, S. J. (2018). Economic burden of neonatal sepsis in sub-Saharan Africa. *BMJ Global Health, 3*(1). https://doi.org/10.1136/bmjgh-2017-000347

Raturi, A., & Chandran, S. (2024). Neonatal Sepsis: Aetiology, Pathophysiology, Diagnostic Advances and Management Strategies. *Clinical Medicine Insights: Pediatrics*, *18*, 11795565241281337. https://doi.org/10.1177/11795565241281337

Raymond, S. L., Stortz, J. A., Mira, J. C., Larson, S. D., Wynn, J. L., & Moldawer, L. L. (2017). Immunological Defects in Neonatal Sepsis and Potential Therapeutic Approaches. *Frontiers in Pediatrics*, *5*. https://doi.org/10.3389/fped.2017.00014

Rubio-Mora, E., Bloise-Sánchez, I., Quiles-Melero, I., Cacho-Calvo, J., & Cendejas-Bueno, E. (2025). Neonatal sepsis: Epidemiology and comparison between preterm and term newborns. *Enfermedades infecciosas y microbiologia clinica (English ed.)*, *43*(4), 197-204.

Salama, B., & Tharwat, E. M. (2023). A case control study of maternal and neonatal risk factors associated with neonatal sepsis. *Journal of Public Health Research*, *12*(1). https://doi.org/10.1177/22799036221150557

Shane, A. L., Sánchez, P. J., & Stoll, B. J. (2017). Neonatal sepsis. *The Lancet*, *390*(10104), 1770–1780. https://doi.org/10.1016/S0140-6736(17)31002-4

Singh, M., Alsaleem, M., & Gray, C. P. (2025). Neonatal Sepsis. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK531478/(Singh et al., 2025)

Sizear, M. I., & Rashid, M. (2024). Urgent need to address increasing caesarean section rates in lower-middle-income countries like Bangladesh. *Frontiers in Global Women's Health*, *5*. https://doi.org/10.3389/fgwh.2024.1365504

Todd, I. M. F., Magnus, M. C., Pedersen, L. H., Burgner, D., & Miller, J. E. (2024). Caesarean section and risk of infection in offspring: systematic review and meta-analysis of observational studies. *BMJ Medicine*, *3*(1), e000995. https://doi.org/10.1136/bmjmed-2024-000995

Utomo, M. T. (2010). Risk Factors of Neonatal Sepsis: A Preliminary Study in Dr.Soetomo Hospital. *Indonesian Journal of Tropical and Infectious Disease*, *1*(1), 23-26. https://doi.org/10.20473/ijtid.v1i1.3718

Wu, M., Deng, Y., Wang, X., He, B., Wei, F., & Zhang, Y. (2024). Development of risk prediction nomogram for neonatal sepsis in Group B Streptococcus-colonized mothers: a retrospective study. *Scientific Reports*, *14*(1), 5629.

Yadav, P., & Yadav, S. K. (2022). Progress in Diagnosis and Treatment of Neonatal Sepsis: A Review Article. *JNMA: Journal of the Nepal Medical Association*, *60*(247), 318–324. https://doi.org/10.31729/jnma.7324

A Cross-Sectional Observational Study of Rural and Urban Differences in Healthcare Access and Emergency Room Utilization in Canada

Afreen Deen, Mina Aldoori, Jules Ciulpan

McMaster University

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Group 11

Dr. Janet Pritchard

April 8th, 2025

Introduction

Canada is one of the many countries worldwide with a universal health system funded by taxes. The Canada Health Act provides provincial funding to ensure that every resident of the country is guaranteed public health insurance that covers medically necessary services, including hospital and physician care (Government of Canada, 2004). The Canada Health Transfer, as outlined in the Canada Health Act, provides federal funding that is then given to provinces (Government of Canada, 2004). While this system is designed to equitably distribute healthcare funding, in practice, accessibility to services is inconsistent, and there are prominent geographic disparities between rural and urban settings (Kirby & Yabroff, 2020). Though there is no consistent definition of urban and rural areas across research, Statistics Canada defines an urban area as a settlement with a population of 1,000 people or a population density of 400 people per square kilometer (Statistics Canada, 2022). Any region that falls short of this criteria is considered rural (Statistics Canada, 2022).

One of the issues faced by rural residents is the limited availability of primary care providers (Matsumoto et al., 2017). Urban emergency departments (EDs) have different service providers and treatment services that are proximal and widely available (Matsumoto et al., 2017). By offering alternative services like walk-in clinics and urgent care centers, ED patients generally have access to a broad network of care even during evenings and weekends (after primary care hours), allowing individuals with less serious injuries or illnesses to receive timely care while keeping ED resources for more critical cases (Matsumoto et al., 2017). Conversely, rural ED environments use the same group of general practitioners (GPs) who provide daytime primary care to staff the ED after hours (Matsumoto et al., 2017). Additionally, rural communities typically lack a discrete set of emergency service providers or alternative after-hours options, making the local ED the default site for urgent care (Matsumoto et al., 2017). High-frequency users are defined as those visiting the ED six or more times. They represent a larger proportion of the patient population in rural areas compared to urban areas. In urban areas, these users represent 3-10% of the urban participants, while in rural areas, they account for over 30% of all ED visits (Matsumoto et al., 2017). This means rural EDs function as an extension of primary care, where even minor injuries like sprains or lacerations require ED visits, unlike urban centers where alternative services exist (Matsumoto et al., 2017). This matters because relying on rural EDs for both emergency and minor care not only strains limited resources, increases wait times, and creates a burden on overworked healthcare providers but also limits access to specialized care, enhancing health disparities between rural and urban populations.

Beyond emergency department use, rural and urban healthcare facilities differ structurally. Urban clinics provide walk-in services and access to specialists and diagnostic labs, which promotes timely and comprehensive care for patients (Haggerty et al., 2007). Rural clinics, however, lack walk-in services and a wide array of professional help, limiting options for immediate care (Haggerty et al., 2007). While 33.3% of urban clinics offer walk-in services during days, evenings, and weekends, only 17.5% of rural clinics provide the same level of accessibility (Haggerty et al., 2007). This impacts how residents in different regions seek medical attention. These examples illustrate the limited alternatives for rural patients, making patients less likely to see their general practitioner compared to their urban counterparts, leading to increased ED visits (Clark et al., 2021).

The overuse of the emergency department not only affects individual patients but increases wait times, straining the Canadian healthcare system (Canadian Agency for Drugs and Technologies in Health, 2023). 17.5% of Canada's population is over 65 years old, and this number is projected to grow to close to 25% of Canada's total population by 2040 (Public Health Agency of Canada, 2020). Older patients and those with more prevalent chronic diseases need continuous access to general practitioners to manage complex medical issues. As of 2024, 8% of Canadians over the age of 45 report not having adequate access to healthcare (Alemu et al., 2024). Without reliable GP access, these populations may struggle to receive timely care and must rely on the emergency department. Improving primary care in rural areas could help reduce emergency department reliance (Clark et al., 2021).

The available literature on the accessibility of primary and emergency care was conducted over the past twenty years and may not reflect the current conditions of rural and urban healthcare centres. Treating rural and urban areas as discrete topics of research has its benefits, yet a comparative analysis is needed to contrast rural-urban differences. Existing literature would be more robust with more reviews. Regional disparities have not been sufficiently examined, particularly in terms of the correlation between physician availability and emergency room usage, in rural and urban areas. The primary objective of this study was to determine if living in rural areas correlates with increased difficulty in accessing a general practitioner over the past year compared to those living in urban areas for Canadians over 18 years old. The secondary objective was to determine if living in rural areas correlates with increased hospital visits for those who cannot get an appointment with a GP compared to those living in urban areas for Canadians over 18 years old.

Methods

Study Design and Database

The database used as the basis for this analysis is an online questionnaire conducted by the Angus Reid Institute, polling randomly selected Canadians within the Angus Reid forum from 2015-09-10 to 2015-09-14. This four-day period captures the respondents' answers leading up to the 2015 federal election. The database was designed to investigate the politics of healthcare in Canada and the intentions and outlooks of prospective voters, focusing on healthcare (Angus Reid Institute, 2016). This database is a cross-sectional and observational survey (Angus Reid Institute, 2016). Questions aimed to capture respondents' voting intentions and perceptions of the healthcare system, based on questions like their ability to book an urgent appointment within a year, as well as demographic data like geographical location (Angus Reid Institute, 2016).

Study Participants

The Angus Reid Institute randomly selected Canadians aged 18 years and older from a population of panel members part of the Angus Reid Forum (2016). The survey's sample size before inclusion and exclusion was 1532 cases, with 149 variables per record (Angus Reid Institute, 2016). The database was sourced from Borealis, which contained a codebook that selected dependent variables pertaining to GP appointments ("Difficulty in booking an urgent appointment with your GP") and ER visits ("Can't get appointment: Go to Hospital ER") as annual measures. Through listwise deletion, records without a response to either of these dependent variables were deleted from the dataset. The sample size for this analysis decreased to 1290, resulting in 242 missing values.

This study compares demographic data from rural and urban participants. These different exposures were reflected in the variables arf_Urban_Rural and arf_Region_Rollup. The latter lists all provinces, groups Atlantic provinces together, and excludes data from the Territories and Indigenous populations on reserves (Angus Reid Institute, 2016). Other geographical variables were excluded (6/8). Rurality was determined as populations below 1000 people or a population density of 400 people/sq km (Statistics Canada, 2021).

The survey had an age variable with three categories: 18-34, 35-54, and 55+. For this study, several variables were excluded: all race (12/12) and birthplace questions (1/1), most demographic variables (11/14), including gender, marital status, children in household, employment, household size, and homeownership. Variables related to previous voting (2/2) and issues/voting intentions (21/21) were also not considered. Within the healthcare questions, any questions assessing specialist use/satisfaction (83/88) were excluded. All excluded variables did not contribute to the outcomes of the study.

Study Outcomes

The primary outcome in this study was to determine the “Difficulty in booking an urgent appointment with your GP.” The response options were ordinal: “Easy - (1-2 days)”, “Medium - (3-4 days)”, and “Difficult - (7+ days)”. This demonstrates that even among Canadians rostered with a family doctor, the availability of urgent care is limited by geographic location (Rahman et al., 2023). In an Ontario-based cross-sectional survey, users resorted to walk-in clinics when same-day or next-day appointments with a GP were unavailable and tended to live in medium-sized or large urban areas with lower incomes (Rahman et al., 2023). This study’s data, validated by past literature, could inform policies to improve primary care access and continuity.The prompt for the secondary research outcome was “Can’t get appointment: Go to Hospital ER (annually)” with response options of “No, have not”, “Once”, “2 to 3 times”, “4 to 5 times”, or “6 or more times” (Angus Reid, 2016).Annual measures allow for international comparisons. For example, the Commonwealth Fund survey of adults in the top ten high-income countries and their access to primary health care (Doty et al., 2021). Additionally, annual reports capture enough data to be meaningful without requiring respondents to recall too much healthcare history (Dalziel et al., 2018).

Additional Data Collection

The geographic variable arf_Urban_Rural was used to split the data presented in Table 1 into urban and rural, and Region (arf_Regional_Rollup in the original dataset) is also a geographic variable of interest to further study provincial differences. Based on a literature search, there are multiple investigations into regional disparities in healthcare in Canada. For example, one study demonstrated that rural areas were less likely to have access to a primary care provider compared to cities close to a census metropolitan area (Shah et al., 2019). The question preceding the primary outcome in the survey was “How often do you typically see your family doctor/GP (for yourself)” in a year, using 2015. Other binary “yes/no” questions like GP and ER visits also supported the dependent variables of interest. Variables like income and education level of participants were measured in Table 1 as well. Studies like Rahman et al.’s 2023 study showed healthcare access varied based on socioeconomic factors.

Statistical Analysis

Data is summarized in Table 1 by outcomes of interest in rows and exposure (rural) and control (urban) groups in columns. The sample size for rural and urban residents was denoted with asterisks to mark missing cases. Outcomes of interest were broken down into respondent options, and then counts were written for each option, followed by the proportion as a percentage of that category's sample size. This study's research question aims to assess if there are group differences between rural and urban residents. Since the data is categorical and cannot be tested for normality, a chi-squared test was used as the primary statistical analysis to measure between-group differences. Spearman's rho correlation was used as the secondary statistical test because it was suitable for ordinal variables, allowing ranking of categorical options based on GP access difficulty or ER visit frequency. This helped determine the directionality of the target variables, ensuring geographic location influences the outcome rather than an inverse relationship. An α-value of 0.05 was used to determine if this difference was statistically significant between rural and urban areas. The software used for data organization and statistical testing was SPSS (version 30.0.0.0).

Results

Participant Recruitment

The starting dataset consisted of 1532 cases and 87 categorical variables (Angus Reid Institution, 2015). This study included 1311 participants and 9 categorical variables. Listwise deletion was conducted if participants did not answer the following questions: "Easy or difficult to get an appointment to see your family doctor/GP (other than booking ahead)" and "Can't get appointment: Go to Hospital ER (annually), which served as the study's primary and secondary dependent variables respectively.

Characteristics of Population

Table 1 describes the demographic characteristics of the 1311 participants from the Angus Reid Forum Panel (Angus Reid Institute, 2015). In urban areas, 40.4% were aged 35-54, and 40.1% were 55+, indicating that the majority were 35 or older. In contrast, a larger proportion of rural participants were aged 55 or older (49.4%). University education or higher was more common in urban areas compared to rural areas (21.2%). In terms of income, 38.6% of rural residents earned less than $50,0000 annually in contrast to their urban counterparts who made up 33.2%.

Table 1. Comparison of Healthcare Access and Hospital Visits Between Rural and Urban Participants in Canada in 2015 [N=1311]

		Rural n=241*	Urban n=1070**
Categorical variables presented as *n* (%)			
Age of Participants	18-34	36 (14.8)	208 (85.2)
	35-54	86 (16.6)	433 (83.4)
	55+	119 (21.7)	429 (78.3)
Education of Participants	High School or less	121 (22.4)	418 (77.6)
	College/Tech School	69 (16.2)	357 (83.8)

	University and above	51 (14.7)	295 (85.3)
Income of Participants	<$50K	93 (20.8)	355 (79.2)
	$50-99K	75 (17.1)	363 (82.9)
	$100K+	41 (16.9)	202 (83.1)
	Don’t know/Refuse to answer	32 (17.6)	150 (82.4)
Region (Provinces)	British Columbia (BC)	30 (16.9)	147 (83.1)
	Alberta	17 (12.8)	116 (87.2)
	Saskatchewan (SK)	18 (20.5)	70 (79.5)
	Manitoba (MB)	13 (14.1)	79 (85.9)
	Ontario (ON)	81 (17.5)	382 (82.5)
	Quebec (QC)	46 (17.2)	221 (82.8)
	Atlantic	36 (39.6)	55 (60.4)

Frequency in seeing your GP	Once a month or more	19 (22.9)	64 (77.1)
	A few times a year	111 (17.2)	536 (82.8)
	Once a year or so	81 (19.5)	335 (80.5)
	Every couple of years	20 (16.0)	105 (84.0)
	Less often than every couple years	10 (25.0)	30 (75.0)
Difficulty in booking an urgent appointment with your GP ***	Easy (1-2 days)	61 (16.5)	309 (83.5)
	Medium (3-4 days)	115 (19.7)	469 (80.3)
	Difficult (7+ days)	65 (18.2)	292 (81.8)
Can't get appointment: Go to Hospital ER (annually)*	No, have not	168 (17.0)	823 (83.0)
	Once	45 (21.7)	162 (78.3)
	2 to 3 times	21 (22.3)	73 (77.7)
	4 to 5 times	5 (50.0)	5 (50.0)

	6 or more times	2 (22.2)	7 (77.8)
Past year: Visited a family doctor or GP	Not Selected	37 (18.8)	160 (81.2)
	Selected	204 (18.3)	910 (81.7)
Past year: Visited a hospital emergency room	Not Selected	170 (16.9)	833 (83.1)
	Selected	71 (23.1)	237 (76.9)

*Missing cases *n* = 36; **missing cases *n* = 185; ***listwise deletion performed on dependent variables

Primary Analysis

Figure 1 illustrates a Chi-Square test highlighting the association between residence and difficulty in getting an appointment with a GP. The Pearson chi-square value was 1.5561, with a degree of freedom of 2 (p=0.458). Figure 2 illustrates a Chi-square test highlighting the association between residence and ER visit frequency when unable to access a GP. The Pearson chi-square value of 10.638 with degree of freedom of 4 (p= 0.031).

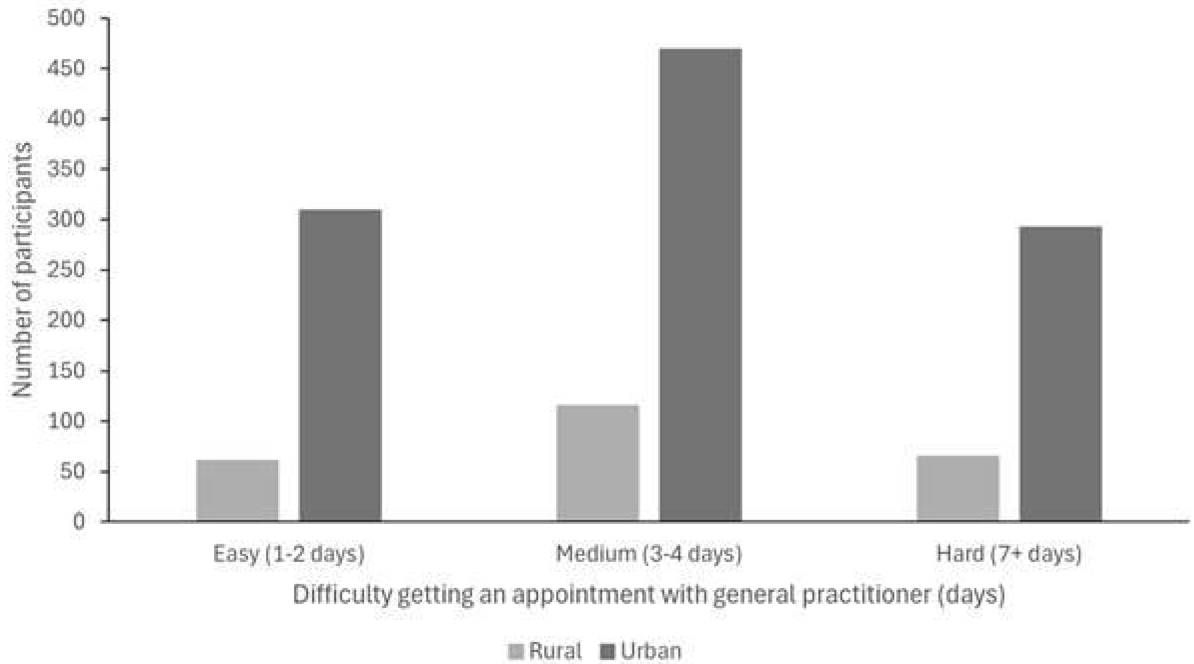


Figure 1. **A comparison of rural and urban participants' responses to difficulty getting an appointment with general practitioners.** This was measured by the number of days it took to book the appointment. This was based on a chi-square analysis. Paired bar graphs demonstrated the number of rural and urban participants that reported easy [N = 61; N = 309], medium [N = 115; N = 469], and hard [N = 65; N = 292] difficulty, respectively. The light bars represent rural participants while the dark bars represent urban participants.

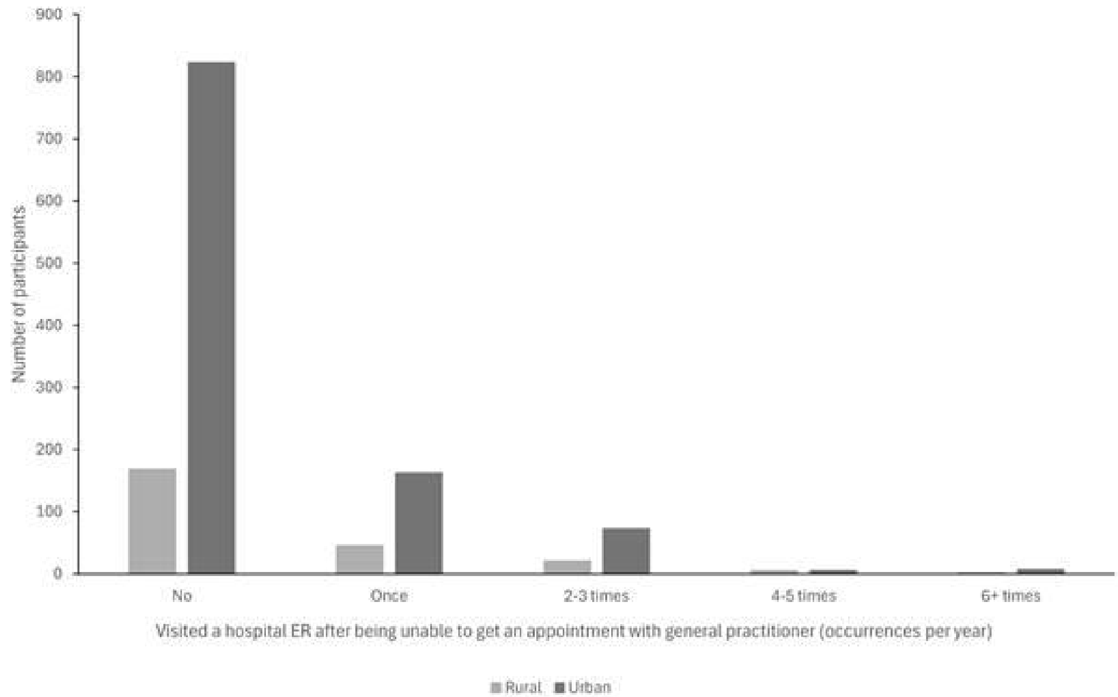


Figure 2. A comparison of rural and urban participants' responses to if a hospital ER was visited after being unable to book an appointment with a general practitioner. This was measured using occurrences within a year. This was based on a chi-square analysis. Paired bar graphs demonstrated the number of rural and urban participants that reported no (0 times) [N = 168; N= 823], once [N = 45; N = 162], 2-3 times [N = 21; N = 73], 4-5 times [N = 5; N = 5], and 6+ times [N = 2; N = 7] visiting a hospital ER, respectively. The light bars represent rural participants while the dark bars represent urban participants.

Secondary Analysis

A Spearman's Rho correlation was used to examine the relationship between difficulty in getting a GP appointment and urban vs. rural residence (Table 2). The correlation coefficient was $\rho = .017$, indicating a very weak positive relationship ($p = .538$). A second Spearman's Rho correlation assessed the relationship between ER visit frequency and residence status. The correlation coefficient was $\rho=.067$, also indicating a very weak positive relationship ($p=.016$).

		Urban vs Rural
Difficulty in getting an appointment with a general practitioner?	Correlation Coefficient	0.017
	Sig. (2-tailed)	0.538

Table 2. Association between urban/rural residency and difficulty accessing GP appointments. Spearman's *rho* correlation results examined the association between participants living in urban and rural environments [N = 1311] and participants' difficulty in getting an appointment with a general practitioner.

		Urban vs Rural
Can't get an appointment: Go to ER?	Correlation Coefficient	0.068
	Sig. (2-tailed)	0.014

Table 3. Association between urban/rural residency and hospital ER use when unable to access GP. Spearman's *rho* correlation results examined the association between participants living in urban and rural environments [N = 1311] and if participants go to hospital ERs if an appointment cannot be secured.

Discussion

Main Findings

The primary aim of this research study was to determine if living in rural areas correlates with an increased difficulty in accessing a general practitioner in the past year in comparison to those living in urban areas within Canadian adults. The secondary aim of the research study was to determine for Canadians over the age of 18, if living in rural areas correlates with increased hospital ER visits for those who cannot get an appointment with a GP in comparison to those living in urban areas. Chi-square value found no statistically significant difference between urban and rural residents in difficulty accessing their GP. The Spearman correlation coefficient illustrated that there was no meaningful relationship between these variables. The data may not have been sufficient to show a significant difference because of the high number of individuals in both groups reporting having a GP. This means that difficulty in access does not pertain to finding a GP, but rather, booking an appointment within a burdened healthcare system (Matsumoto et al., 2017). For the secondary objective, there was a statistically significant chi-square value, demonstrating the difference between rural and urban residents when it comes to visiting the ER when unable to get a GP appointment. The Spearman test found a very weak, yet statistically significant correlation between rural areas and ER usage frequency. This suggests that rural residents are more reliant on the ER when primary care is inaccessible, even though difficulty in accessing the GP by appointment did not differ by location.

Comparison to Other Studies

While this study did not find a difference between urban and rural GP access, Clark et al. found a statistically significant difference between groups after adjusting for confounding variables such as income, education, and chronic illness through logistic regression, with rural residents less likely to see a GP (2021). The differences in results could be because this study collected data based on perceived difficulty accessing GPs, while Clark et al. (2021) counted the number of GP visits. This could mean that rural residents may not perceive less visits as a limiting factor in accessibility. Though, it is important to consider that Clark et al. (2021) only measured rurality by count, as opposed to other measurement methods, like population density and distance to urban centers, which could have influenced the results. Therefore, this study contributes to showing the relationship between being a rural resident and GP access in a new perspective with a different definition of rurality. For the secondary outcome, this study found a difference between urban and rural residents in their reported increase in hospital ER visits after inability to book a GP appointment. In a 2007 study by Ionescu-Ittu et al., which focused on elderly populations, they found that rurality was a significant predictor of increased ER use. Both studies had a predominantly older population, where Ionescu-Ittu et al. (2007) focused on individuals over 65 and this study with a rural majority in ages over 55. Though, it is important to consider that Ionescu-Ittu et al. (2007) relied solely on Quebec administrative health databases, which lacks generalizability and information on patients' perceived health status and reasons for visiting the ER, which could be confounding factors that influence ER use. Therefore, this study contributes to revealing that perceived difficulty in accessing a GP is a meaningful predictor of increased ER use, especially among older rural residents, emphasizing the importance of patient experiences in addressing healthcare access gaps.

Limitations

A limitation to this dataset is that missing cases, which had to be accounted for in the statistical analysis, have the potential to impact the results of the study, as certain participants had to be excluded for not providing answers. Additionally, in a chi-square analysis, when a contingency table of 3 x 2 or greater is used, the results may not be reliable as it becomes difficult to assess statistical significance. The proportions may differ, as there is only one chi square value for the entire figure (Norman, 2014). Finally, the dataset excluded certain populations, such as participants living in the territories of Canada and people living on reserves, who may have differing healthcare utilization patterns which lessens the generalizability of this study. Another limitation is that each study uses different terms and rural classification criteria, and these discrepancies may explain the lack of consensus on measuring accessibility.

Implications and Future Directions

Some implications are that the findings of this study can inform federal policy makers on the healthcare access disparities between urban and rural populations. Highlighting the need for resource allocation for areas that lack primary care and have overburdened emergency rooms. A key future direction for this study is linking virtual care initiatives to improve access and reduce avoidable ER visits for underserved areas. Additionally, expanding the study to include excluded populations such as Indigenous reserve populations and those living in Territories, while also making the study longitudinal would provide a more comprehensive understanding of healthcare access and its impact across diverse groups.

References

Alemu, F. W., Nicholson, K., Wilk, P., Thornton, J. S., & Ali, S. (2024). Unmet need and access to family physicians: A national analysis using the Canadian Longitudinal Study on Aging. *SSM - Health Systems, 14*, 100028. https://doi.org/10.1016/j.ssmhs.2024.100028

Angus Reid Institute. (2016). *Election 2015: The politics of health care [Canada]* (Version 3.2) [Data set and codebook]. Borealis. https://doi.org/10.5683/SP3/JDA1CO

Clark, K., John, P. S., Menec, V., Cloutier, D., Newall, N., O'Connell, M., & Tate, R. (2021). Healthcare utilisation among Canadian adults in rural and urban areas – The Canadian Longitudinal Study on Aging. *Canadian Journal of Rural Medicine, 26*(2), 69. https://doi.org/10.4103/CJRM.CJRM_43_20

Dalziel, K., Li, J., Scott, A., & Clarke, P. (2018). Accuracy of patient recall for self-reported doctor visits: Is shorter recall better? *Health Economics, 27*(11), 1684–1698. https://doi.org/10.1002/hec.3794

Doty, M. M., Tikkanen, R. S., FitzGerald, M., Fields, K., & Williams, R. D. (2021). Income-related inequality in affordability and access to primary care in eleven high-income countries. *Health Affairs, 40*(1), 113–120. https://doi.org/10.1377/hlthaff.2020.01566

Government of Canada. (2004, July 26). *Canada Health Act* [Navigation page]. https://www.canada.ca/en/health-canada/services/health-care-system/canada-health-care-system-medicare/canada-health-act.html

Ionescu-Ittu, R., McCusker, J., Ciampi, A., Vadeboncoeur, A.-M., Roberge, D., Larouche, D., Verdon, J., & Pineault, R. (2007). Continuity of primary care and emergency department utilization among elderly people. *CMAJ, 177*(11), 1362–1368. https://doi.org/10.1503/cmaj.061615

Matsumoto, C., O'Driscoll, T., Madden, S., Blakelock, B., Lawrance, J., & Kelly, L. (2017). Defining "high-frequency" emergency department use. *Canadian Family Physician, 63*(9), e395–e399. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5597030/

Norman, G. R. (2014). *Biostatistics: The bare essentials* (4th ed.). People's Medical Publishing

House. Public Health Agency of Canada. (2020). *Aging and chronic diseases: A profile of Canadian seniors* [Education and awareness]. https://www.canada.ca/en/public-health/services/publications/diseases-conditions/aging-chronic-diseases-profile-canadian-seniors-report.html

Rahman, B., Costa, A. P., Gayowsky, A., Rahim, A., Kiran, T., Ivers, N., Price, D., Jones, A., & Lapointe-Shaw, L. (2023). The association between patients' timely access to their usual primary care physician and use of walk-in clinics in Ontario, Canada: A cross-sectional study. *CMAJ Open*, *11*(5), E847–E858. https://doi.org/10.9778/cmajo.20220231

Shah, T. I., Clark, A. F., Seabrook, J. A., Sibbald, S., & Gilliland, J. A. (2020). Geographic accessibility to primary care providers: Comparing rural and urban areas in Southwestern Ontario. *Canadian Geographer / Géographe Canadien*, *64*(1), 65–78. https://doi.org/10.1111/cag.12557

Statistics Canada. (2021, November 17). *Dictionary, Census of Population, 2021 – Rural area (RA)*. https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/az/Definition-eng.cfm?ID=geo042

Wilson, C. R., Rourke, J., Oandasan, I. F., & Bosco, C. (2020). Progress made on access to rural health care in Canada. *Canadian Family Physician*, *66*(1), 31–36. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7012120/

Depressive Symptoms and Substance Use in Canadian Adolescents Aged 10–15: A Cross-Sectional Analysis

Ajandan Arulananthan, Daniel Ravendra, Vengadesh Baskaran

Department of Life Sciences, McMaster University

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8, 2025

Introduction

The prevalence of substance abuse during adolescence is a serious public health concern due to the plethora of implications it has on physical, social, psychological, and academic functioning. Characterized by rapid physiological and psychological growth, adolescence is a phase where individuals are increasingly susceptible to engaging in risky activities, such as substance abuse (McKowen et al., 2013). The repercussions of substance use during adolescence are profound, encapsulating not only immediate risks such as injuries, accidents, and violence, but also long-term implications including mental health disorders, social isolation, chronic health conditions, and reduced economic opportunities (McKowen et al., 2013). An increasingly relevant field of study highlights the necessity of identifying and understanding underlying factors and influences that predispose adolescents to substance abuse in order to integrate and promote effective prevention strategies (McKelvey et al., 2017).

Furthermore, depression, which encompasses a significant psychological disorder among adolescents, has a profound effect on their overall quality of life and development. Recently, the prevalence of depression among adolescents has been increasingly high, with recent epidemiological studies reporting that approximately 17% of adolescents experience major depressive episodes in North America (Ghandour et al., 2019). Among adolescents, depression is commonly associated with feelings of sadness, hopelessness, and low self-worth, often resulting in social withdrawal and reduced academic performance. Moreover, depressive disorders have been extensively linked to increased engagement in substance use behaviors, as adolescents may cope with their negative emotions through alcohol, marijuana, or tobacco use (Lai et al., 2015).

Existing research supports a strong correlation between depression and substance use among adolescents, indicating that depressive symptoms significantly elevate the risk of tobacco, alcohol, and drug use (Dyer et al., 2019). For example, adolescents experiencing depressive episodes tend to exhibit higher tendencies to engage in illicit drug use and binge drinking compared to their non-depressed counterparts (McKowen et al., 2013). However, while the correlation is well-documented, the precise nature of this relationship remains underdeveloped, particularly regarding the distinctions between adolescents with clinical diagnoses and those experiencing subclinical or emerging depressive symptoms.

Previous studies examining this relationship often focus on general associations and predominantly utilize cross-sectional designs, limiting their capacity to infer temporal patterns or potential risk trajectories. Research by Dyer et al. (2019), for instance, identified increased alcohol use among adolescents diagnosed with depression, yet failed to explore the underlying mechanisms and directionality of this association. Moreover, many studies rely heavily on self-reported data, which introduces risks of recall bias and underreporting due to social stigma (Lai et al., 2015). The lack of demographic differentiation—such as age, sex, socioeconomic status, and family context—further limits the generalizability and practical application of findings (McKowen et al., 2013; Lai et al., 2015).

To address these limitations, the present study adopts a longitudinal design utilizing data from the National Longitudinal Survey of Children and Youth (NLSCY) to assess how varying levels of depressive symptoms over time are associated with patterns of substance use in Canadian adolescents aged 10 to 15. The study focuses on repeated measures of both depressive symptoms and self-reported use of alcohol, marijuana, and tobacco over multiple developmental periods. This design enables the exploration of developmental pathways and trajectories, incorporating key demographic variables such as sex, socioeconomic status, and parental support, with the aim of informing targeted, early interventions.

To effectively explore this relationship, the present study is structured around the PICO framework, a model that enhances clarity in research design by narrowing the scope and specifying key elements. The population of interest includes Canadian adolescents between the ages of 10 and 15, a developmental period marked by significant psychological vulnerability and exposure to external stressors. The primary exposure being examined is the presence and severity of depressive symptoms, which will be assessed through standardized psychological evaluations administered across multiple waves of data collection. These adolescents will be compared to their non-depressed or minimally depressed counterparts, serving as the comparison group. The central outcome measured is the self-reported use of substances, specifically alcohol, marijuana, and tobacco across the same time period.

Accordingly, this study is guided by the following research question: Among Canadian adolescents aged 10 to 15, how are depressive symptoms associated with substance use behaviors, including alcohol, marijuana, and tobacco use, compared to adolescents reporting lower or no depressive symptoms? Through the use of longitudinal data, this study seeks to track changes and patterns over time, offering insights into how mental health trajectories intersect with substance use behaviors during a formative stage of development. By examining these temporal associations alongside key demographic variables, the findings aim to support the design of evidence-based prevention strategies tailored to the complex realities faced by adolescents experiencing depressive symptoms.

Methods

Strong Methods form the backbone for the credibility of any empirical study, especially in disciplines examining multifactorial human behaviors like mental health and substance use among adolescents. In our research, the approach taken to collect data, select participants, measure outcomes, and analyze statistical significance was strategically planned to yield valid and interoperable insights. This section outlines the methodological framework of our study, evaluating its suitability in answering the central research question: Does depressive symptoms increase the likelihood of substance use among Canadian adolescents?

Study Design

This study utilizes a cross-sectional observational design, drawing data from the *National Longitudinal Survey of Children and Youth, 1998-1999 [Canada]: Cycle 3, 10-15 Year Olds* (NLSCY), A nationally representative database maintained by Statistics Canada in collaboration with Human Resources Development Canada. Although the NLSCY was designed as a longitudinal resource capturing a plethora of developmental milestones, for this phase of the study, Data was treated as cross-sectional, selecting a single time-point to explore contemporaneous associations. This design choice aligns with the study's immediate objective: to identify and quantify associations between depressive symptoms and substance use behaviors in adolescents. While this sacrifices the ability to make causal inferences, it provides a strong foundation for establishing statistically significant associations that can inform longitudinal follow-up studies.

Study Participants

The target population consisted of adolescents aged 10 to 15 years, identified through the age grouping variable. Inclusion criteria mandated complete responses to both depression and substance use variables. The focus on this age range was intentional, aiming to capture the developmental window where both depression and initial substance experimentation are most likely to co-occur. Furthermore, participants were excluded if they fell outside the age range or if key variables were missing, ensuring the integrity of statistical inference.

Operationalization of Key Variables

Substance use outcome was derived from 3 dichotomous variables, including tobacco, alcohol, and marijuana. This multifaceted approach transcends the limitation of many existing studies that isolate single substances, thereby offering a more ecologically valid measure of adolescent substance engagement. Depression was measured through multiple self-report indicators, allowing for sensitivity analyses across operational definitions. This reflects both symptom frequency and perceived distress. Although self-reported data is inherently vulnerable to social desirability and recall biases, its use is well-supported in adolescent research, particularly when anonymity is ensured.

Confounding Variables

In any observational study that examines physiological and behavioral outcomes, there is a significant threat to internal validity that is caused by confounding variables. This study recognizes this threat and carefully accounts for a range of covariates that are linked to both adolescent substance use and depressive symptoms. These variables were selected not only for their relevance but also for their availability and quality within the NLSCY dataset. Sex and age were included due to known differences in development. Socioeconomic status addressed influences such as access to mental health care and environmental stressors. Additionally, social support, defined through the “Friends Score,” was analyzed as both a confounder and potential protective factor. Finally, anxiety disorder indicators were also tested to address comorbidity, accounting for misattribution of variance to depression alone.

Statistical Analysis and Justification

All analyses were conducted using IBM SPSS Statistics. We began with descriptive statistics, summarizing demographic and psychosocial characteristics. Continuous variables, normality was assessed using the Shapiro-Wilk test, and categorical variables through frequencies. Chi-square tests evaluated bivariate associations between depression indicators and each substance use variable, followed by binary regression to account for confounders. Moreover, The threshold for statistical significance was set at $p < 0.05$, to allow for a balance between minimizing Type I and Type II errors in observational studies.These models enabled us to isolate the relationship between depressive symptoms and substance use.

Appropriateness of the Methodology

Overall , the analytical approach is well-aligned with the research question. The use of the NSCLY dataset, paired with the statistical control for confounding, allows for a nuanced understanding of the relationship between depression and substance use in adolescents. Even though this study is cross-sectional by design, the longitudinal origins of the dataset allow for future tracking. The methods used in this study are appropriate not only from a statistical standpoint but also from an ethical perspective. Through outlining the rationale clearly for analysis choices and maintaining the norms of epidemiology testing, this study sets a strong foundation for future research and policy engagement.

Results

Table 1:

Descriptive Characteristics of Study Participants (n = 5267)

Variables		**Frequencies**
Number of Participants		5267
Sex	Female	2659 (50.5)
	Male	2608 (49.5)
Age	10-11 year olds	1609 (30.5)
	12-13 year olds	1832 (34.8)
	14-15	1826 (34.7)
Ever had a drink of alcohol	Yes, at least one drink	1638 (31.1)
	I have only had a few sips	1299 (24.7)
	No	2330 (44.2)
Region of Canada	Atlantic	1133 (21.5)
	Quebec	825 (15.7)
	Ontario	1250 (23.7)
	Parries	1278 (24.3)
	British Colombia	415 (7.9)
	Not states	366 (6.9)
Have you ever smoked	Yes	1750 (33.2)
	No	3336 (63.3)
	Prefer not to answer	181 (3.4)
Ever tried Marijuana/hash	Yes	575 (10.9)
.	No	2755 (52.3)
	I don't know what that drug is	107 (2.3)
Emotional Disorder Score		4744 (90.1)

This table presents key demographic and behavioral characteristics of the adolescent sample. Participants were nearly evenly distributed by gender (50.5% female, 49.5% male) and across three age categories: 10–11 (30.5%), 12–13 (34.8%), and 14–15 (34.7%) years. The majority of respondents reported never smoking (63.3%) or trying marijuana (52.3%), while 44.2% had never consumed alcohol. Emotional Disorder Scores were available for 90.1% of the sample. These figures provide context for the analyses linking depressive symptoms with substance use and other psychosocial factors.

Sample Characteristics

The sample for this analysis was drawn from the National Longitudinal Survey of Children and Youth (NLSCY), which includes responses from 5,539 adolescents. In total, three distinct substances were assessed: marijuana or hash, cigarette smoking, and alcohol consumption. Depressive symptoms were quantified through the Emotional Disorder Score, ranging from 0 to 16. The statistical tests used in the analysis primarily focused on Chi-square tests to evaluate the associations between depressive symptoms and substance use behaviors. These associations were further examined across gender groups.

The sample comprised a broad representation of adolescents between the ages of 10 to 15, with no exclusions based on other demographic characteristics such as socioeconomic status or family structure. The data included self-reported measures of depressive symptoms and substance use, which were crucial for analyzing potential associations between these variables. As expected, the sample showed a diverse range of depressive symptom scores, with a skew toward lower scores, indicating that a majority of adolescents reported mild or no symptoms of emotional disorder.

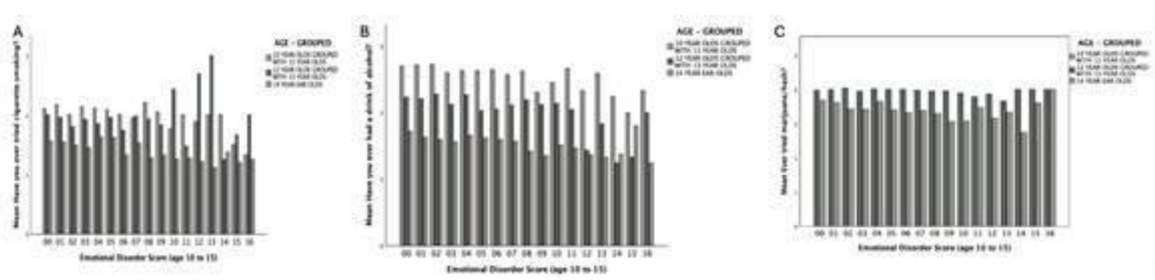


Figure 1. Age-stratified associations between Emotional Disorder Scores and substance use behaviors.

Panel A depicts the mean response to the question “Have you ever tried cigarette smoking?” across Emotional Disorder Scores (0–16), grouped by age. Panel B illustrates the same relationship for alcohol use, and Panel C for marijuana or hash use. In all three panels, older adolescents (particularly those aged 13–14) reported higher levels of substance use at elevated Emotional Disorder Scores. These patterns suggest that age may moderate the association between emotional distress and substance use behaviors, with stronger associations emerging among older youth.

A

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	65.641[a]	32	<.001
Likelihood Ratio	65.336	32	<.001
Linear-by-Linear Association	25.625	1	<.001
N of Valid Cases	4744		

B

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	66.516[a]	32	<.001
Likelihood Ratio	66.429	32	<.001
Linear-by-Linear Association	3.710	1	.054
N of Valid Cases	4744		

C

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	61.200[a]	32	.001
Likelihood Ratio	61.243	32	.001
Linear-by-Linear Association	27.722	1	<.001
N of Valid Cases	3212		

D

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1.044[a]	2	.593
Likelihood Ratio	1.045	2	.593
N of Valid Cases	3437		

E

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	6.303[a]	2	.043
Likelihood Ratio	6.306	2	.043
N of Valid Cases	5267		

F

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	12.379[a]	2	.002
Likelihood Ratio	12.385	2	.002
N of Valid Cases	5267		

Figure 2. Chi-Square Test Results.

Chi-square tests were conducted to examine the associations between Emotional Disorder Scores and three substance use behaviors: alcohol use, cigarette smoking, and marijuana/hash use. The relationship between Emotional Disorder Score and alcohol use (Figure A) was statistically significant, $\chi^2(32) = 65.64$, $p < .001$, indicating that adolescents with higher emotional disorder scores were more likely to report having consumed alcohol. A similar association was observed for cigarette smoking (Figure B), $\chi^2(32) = 66.52$, $p < .001$, though the linear-by-linear association approached significance, $p = .054$, suggesting a weaker trend. Marijuana/hash use (Figure C) also demonstrated a significant association with emotional disorder scores, $\chi^2(32) = 61.20$, $p = .001$, with a strong linear-by-linear trend ($p < .001$). These results provide further evidence of the robust association between elevated depressive symptoms and increased likelihood of substance use in adolescents.

Chi-square tests were performed to assess gender differences in self-reported substance use behaviors among adolescents. The association between gender and alcohol use (Figure D) was not statistically significant, $\chi^2(2) = 1.044$, $p = .593$, indicating no meaningful difference between male and female adolescents in their likelihood of having consumed alcohol. However, the test for gender and marijuana/hash use (Figure E) revealed a statistically significant association, $\chi^2(2) = 6.303$, $p = .043$, suggesting that marijuana use may differ slightly between male and female adolescents. A stronger gender difference emerged in cigarette smoking behaviors (Figure F), with a significant Chi-square result, $\chi^2(2) = 12.379$, $p = .002$, indicating that cigarette use was more strongly associated with gender. These findings suggest that while alcohol use patterns are similar across genders, marijuana and especially cigarette smoking show notable gender-based differences in adolescent substance use.

Tests of Between-Subjects Effects

Dependent Variable: Emotional Disorder Score (age 10 to 15)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	4949.438[a]	5	989.888	123.470	<.001	.135
Intercept	1111.193	1	1111.193	138.601	<.001	.034
CGEHD03B	16.536	1	16.536	2.063	.151	.001
CMMCQ01B	205.315	1	205.315	25.609	<.001	.006
CPMCbS2B	2067.440	1	2067.440	257.875	<.001	.061
CPMCcQ1P	179.945	1	179.945	22.445	<.001	.006
CMMCQ02	1066.786	1	1066.786	133.062	<.001	.033
Error	31716.119	3956	8.017			
Total	93062.000	3962				
Corrected Total	36665.558	3961				

Figure 3: **ANCOVA Results Showing the Effects of Gender and Parenting Variables on Emotional Disorder Scores**

An Analysis of Covariance (ANCOVA) was conducted to examine whether depressive symptoms (measured by the Emotional Disorder Score) differed between male and female adolescents after controlling for family and social factors, including parental support, communication, and socioeconomic status. Results showed that girls reported significantly higher levels of depressive symptoms ($M = 4.24$, $SD = 3.09$) compared to boys ($M = 3.29$, $SD = 2.91$), even when these additional variables were considered. Parental communication and monitoring were strong predictors of lower emotional distress, while parental encouragement was also linked to lower depressive symptoms. Socioeconomic status, however, was not a significant factor once other influences were controlled. The model accounted for approximately 13.5% of the variance in depressive symptoms, suggesting that gender and family dynamics are important contributors to adolescent emotional well-being.

Discussion

This study aimed to investigate the association between depressive symptoms and substance use behaviors, specifically marijuana, cigarette smoking, and alcohol use, among Canadian adolescents aged 10 to 15. Based on analyses from 5,539 participants in the NLSCY, we observed significant associations between higher Emotional Disorder Scores and substance use in the overall sample. Chi-square analyses showed that adolescents with elevated depressive symptoms were more likely to report having tried marijuana ($\chi^2 = 61.20$, $p = 0.001$), alcohol ($\chi^2 = 65.64$, $p < 0.001$), and cigarettes ($\chi^2 = 66.52$, $p < 0.001$), although the linear trend for smoking approached, but did not reach, significance ($p = 0.054$). Age-stratified trends revealed that these associations were more pronounced in older adolescents, particularly those aged 13–14. This suggests that as adolescents mature, the link between emotional distress and substance use may strengthen.

When stratified by gender, depressive symptoms were significantly associated with substance use behaviors only among female participants. Girls with higher Emotional Disorder Scores were significantly more likely to report use of marijuana ($\chi^2 = 87.835$), cigarettes ($\chi^2 = 114.592$), and alcohol ($\chi^2 = 95.973$), all with p-values < 0.001. In contrast, these associations were not statistically significant among male adolescents, with all p-values exceeding 0.05. This gender-specific pattern reinforces the idea that depressive symptoms may play a stronger role in influencing substance use behaviors in females compared to males. An additional ANCOVA revealed that girls reported significantly higher depressive symptoms than boys (M = 4.24 vs.3.29), even after controlling for variables such as parental support and socioeconomic status. Family dynamics, particularly parental communication, encouragement, and monitoring, were all associated with lower levels of emotional distress, whereas socioeconomic status had no significant effect when other variables were held constant.

Comparison to Current Literature

There is a well-established body of literature documenting the link between depressive symptoms and substance use during adolescence. Prior studies have shown that youth experiencing emotional difficulties are at increased risk of initiating or increasing their use of substances such as alcohol, marijuana, and tobacco. The findings from our study reinforces this understanding, but also expands on it by highlighting a clear gender-specific pattern. This distinguishes our findings from some of the broader claims made in existing research.

One of the most relevant studies for comparison is Wilkinson et al. (2016), which investigated the bidirectional relationship between depressive symptoms and substance use in adolescents, with an emphasis on sex differences. Wilkinson’s study found that depressive symptoms predicted marijuana use in males and cigarette smoking in females. Our findings partially agree with WIlkinson's study, as we both observed a statistically significant association between depression and cigarette use in females. However, our studies differ when it comes to males as we found no significant associations across any substances. The absence of significant findings amongst boys contrasts some of the general claims made in literature, and is particularly noteworthy as it adds to this conversation by emphasizing that the association between depression and substance use is not uniform across sexes and should not be treated as such in research or practice.

Strengths and Limitations

This study possesses several methodological strengths that increase the credibility and relevance of our findings. First, we utilized data from the NLSCY, a large, nationally representative dataset that includes detailed demographic, behavioral, and mental health information on Canadian youth. The size and diversity of the sample allow for greater generalizability to the Canadian adolescent population and reduce the likelihood that findings are due to sampling error or population-specific effects. Our study also looked at three forms of substance abuse which included cigarettes, alcohol, and marijuana. This multifaceted approach provides a more accurate analysis of how depressive symptoms relate to substance use behaviors in early adolescence. Lastly, we included several important control variables in our regression models, such as anxiety symptoms, peer support, and socioeconomic status. These confounding factors are known to influence both depression and substance use and their inclusion helps strengthen the internal validity of our results.

While our study provides meaningful insight into the relationship between depressive symptoms and substance use among Canadian adolescents, several important limitations must be considered. Other limitations include our analysis being cross-sectional. As a result, we cannot determine causal direction, whether emotional difficulties lead to substance use, or if substance use contributes to worsening mental health. Second, both depression and substance use were assessed using self-report measures, which are susceptible to recall bias and social desirability bias. Adolescents may underreport behaviors like smoking or drug use, particularly if they perceive these behaviors as stigmatized. Lastly, although we included gender as a stratified variable, we did not account for other important confounders like peer influence, family structure, or access to mental health support. These factors could influence both depression and substance use independently.

Implications

Our findings contribute to the growing body of research on adolescent mental health and substance use, particularly by emphasizing gender-specific risk profiles. The association between depression and substance use in adolescent girls, and its absence in boys, highlights the need for gender-responsive mental health screening and early interventions. These findings suggest that schools, healthcare providers, and community programs should prioritize emotional health screening for girls in early adolescence. Prevention programs could be more effective if they integrate emotional coping strategies and substance use education in a way that is sensitive to sex-based psychological differences. Our study identifies a critical research gap in understanding how adolescent boys experience and respond to depressive symptoms. Future studies should explore alternate behavioral expressions of distress in males, such as aggression or social withdrawal, and investigate whether these behaviors are related to substance use patterns. These findings hold important implications for supporting early, tailored approaches to reducing substance use and improving adolescent mental health outcomes.

Conclusion

In summary, this study examines the relationship between depressive symptoms and substance use among Canadian adolescents aged 10 to 15. We found that emotional disorder scores were significantly associated with the use of alcohol, marijuana, and cigarettes, but only among female participants. These findings contribute to current literature showing a strong link between emotional distress and early substance use in adolescent girls, while also identifying a gap in understanding how this relationship manifests in boys. By identifying clear gender differences in how depressive symptoms relate to substance use, our study highlights the importance of early, tailored mental health interventions. Our results not only address the original research question, but also expand the conversation around gendered risk, offering valuable insight for clinicians, educators, and policymakers working to reduce the burden of youth substance use and mental health challenges in Canada.

References

Dyer, M. L., Heron, J., Hickman, M., & Munafò, M. R. (2019). Alcohol use in late adolescence and early adulthood: The role of generalized anxiety disorder and drinking to cope motives. *Drug and Alcohol Dependence, 204, Article 107480*. https://doi.org/10.1016/j.drugalcdep.2019.04.044

Ghandour, R. M., Sherman, L. J., Vladutiu, C. J., Ali, M. M., Lynch, S. E., Bitsko, R. H., & Blumberg, S. J. (2019). Prevalence and treatment of depression, anxiety, and conduct problems in US children. *The Journal of Pediatrics*, *206,* 256-267. https://doi.org/10.1016/j.jpeds.2018.09.021

Lai, H. M., Cleary, M., Sitharthan, T., & Hunt, G. E. (2015). Prevalence of comorbid substance use, anxiety, and mood disorders in epidemiological surveys, 1990-2014: A systematic review and meta-analysis. *Drug and Alcohol Dependence*, *154,* 1-13. https://doi.org/10.1016/j.drugalcdep.2015.05.031

McKelvey, K. L., Ramo, D. E., & Delucchi, K. (2017). The association between mental health symptoms and use of cigarettes and electronic cigarettes among adolescents. *Drug and Alcohol Dependence*, *179*, 161-166. https://doi.org/10.1016/j.drugalcdep.2017.06.027

McKowen, J. W., Tompson, M. C., Brown, T. A., & Asarnow, J. R. (2013). Longitudinal associations between depression and problematic substance use in the Youth Partners In Care study. *Journal of Clinical Child & Adolescent Psychology*, *42(5),* 669–680.

https://doi.org/10.1080/15374416.2012.759226

Merianos, A. L., Rosen, B. L., Montgomery, L., Barry, A. E., & Smith, M. L. (2018). Impact of perceived risk and depressive symptoms on adolescent marijuana use. *Substance Use & Misuse*, *53(5),* 798-808. https://doi.org/10.1080/10826084.2017.1387567

Wilkinson, A. L., Halpern, C. T., & Herring, A. H. (2016). Directions of the relationship between substance use and depressive symptoms from adolescence to young adulthood. *Addictive Behaviors*, *60*, 64–70. https://doi.org/10.1016/j.addbeh.2016.03.036

Association Between Regular Primary Care Access and Self-Perceived Mental Health in Canadian Emerging Adults

Group 13

Erlin Espina and Evelyn Wang

School of Interdisciplinary Science, McMaster University

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 11, 2025

Introduction

With almost one in five Canadians estimated to have a mental health condition annually (Canadian Institute for Health Information [CIHI], 2023; Statistique Canada, 2023), mental health remains a major public health concern for Canada. Widely acknowledged as the first point of contact for mental health support, primary healthcare is fundamental in early intervention, treatment coordination, and long-term disease management of mental health conditions, including depression and anxiety (World Health Organization [WHO], 2018). In a study published by Flood and colleagues in 2023, many Canadians lack a regular family physician and deal with extended wait times in primary care, which are factors that lead to delayed diagnosis and unbalanced treatment of mental health problems. For example, about 75% of Canadians look to their primary care providers for mental health concerns (Moroz et al., 2020). Rising rates of anxiety and depression combined with reports of unmet mental health treatment requirements are routinely shown by national surveys, including the annual Canadian Community Health Survey (CIHI, 2023; Statistique Canada, 2023). Through cooperative and interdisciplinary approaches, integrated primary health care models, including mental health services within general medical care practices demonstrate the ability to enhance outcomes (Kates, 2009; WHO, 2018). Though primary health care plays a vital role in meeting mental health needs, major structural and systematic obstacles still exist. These include long wait times, geographic inequalities, and provider shortages , challenges further amplified by socioeconomic elements, including cost and stigma (CIHI, 2023; Langlois et al., 2018). Early indicators of depression and anxiety may go unnoticed or untreated when patients are unable to get timely, ongoing, and culturally sensitive primary health care, hence starting a cycle of increasing mental health problems (Galambos et al., 2022; Flood et al., 2023). While several studies have linked poor primary health care access to increased rates of untreated mental health disorders, Canadian-specific evidence, especially those utilizing nationally representative data, is still inadequate.

Extensive research has investigated the link between primary care access and mental health, demonstrating that access to primary health care has important implications on mental health outcomes. Many studies focus on the adult or adolescent population, but fewer have examined emerging adults, which encompasses the ages of 18 to 29, a demographic experiencing significant shifts in life marked with changes in their mental health (Arnett et al., 2014). This new stage in life is a transitory period in which emerging adults experience incredible instability and stark changes coming from adolescence whilst being on the cusp of full adulthood (Arnett et al., 2014). This age group also has the highest prevalence of psychiatric disorders, with over 40% of US young adults aged 18–29 experiencing a psychiatric condition (Kessler, 2005). While the current literature acknowledges the high prevalence of psychiatric disorders in this age group, there is a gap in understanding how primary health care access uniquely affects mental health outcomes in emerging adults.

Our study aims to fill this research gap by investigating the research question: "Among emerging adults aged 18–24 in Canada, how does accessing primary health care compared to not accessing care relate to self-perceived mental health?" To explore this, we will use data collected from the Canadian Community Health Survey from 2017 to 2018 to evaluate a potential association. Access to primary health care will be collected by answering "yes" or "no" to the question "Do you have access to a primary health care provider?" and self-reported mental health will be assessed through the selection of "poor," "fair," "good," and "excellent." Based on the information we collect through this study, we aim to not only fill this gap in existing literature but also to provide evidence that could guide policy improvements to cater to emerging adults' primary healthcare and mental health needs.

Methods

Study Design and Database

To answer the research question, we designed a cross-sectional study that uses data extracted from the 2017-2018 Canadian Community Health Survey (CCHS), which was accessed through the data exploration platform Odesi. The survey reports on information relevant to the health status of Canadians, including health care utilization and determinants of health. The CCHS was designed to supply data for health surveillance and population health research studies. Disseminated data from the CCHS is used to inform health policy and to monitor, plan, and assess program implementations in the context of health. The survey is released every two years and requires data collected from a sample of 130,000 respondents. Respondents are chosen using a multi-stage sampling strategy, sending the survey to provinces and health regions based on population size.

Study Participants

From the original dataset containing 109,470 participants, the final analytic sample included 6819 emerging adults aged 18 to 24, drawn from the 2017–2018 CCHS. This study focused on access to a regular primary healthcare provider being the primary independent variable. The primary dependent variables were mental health statements, specifically a self-perceived mental health state. These primary variables thus informed the inclusion criteria used to select participants from the original database.

Participants were included if they were aged 18–24 years and had complete responses for the primary variables of interest. Participants with non-response answers, including "valid skip," "don't know," or "refusal," to the questions: "Do you have a regular healthcare provider? By this, we mean one health professional that you regularly see or talk to when you need care or advice for your health", and "In general, would you say your mental health is…?" were deleted as these responses were considered to be missing cases. This exclusion was crucial in minimizing missing data bias while maintaining a nationally representative sample, which was consistent with the study's goal of exploring the connections between access to primary healthcare and mental health outcomes during this important transitional phase of life.

Study Outcomes

Primary Outcomes

The availability of a consistent primary health care provider is the main independent variable. Participants were asked to indicate if they had a regular primary healthcare provider by responding to the question: "Do you have a regular healthcare provider? By this, we mean one health professional that you regularly see or talk to when you need care or advice for your health." Researchers have indicated that having a consistent primary healthcare provider can help in the early identification and management of mental health problems. Flood et al. (2023) highlight the importance of continuity in primary care in Canada, pointing out that patients who maintain an ongoing relationship with a primary health care provider tend to experience shorter wait times and benefit from more coordinated care. This ongoing relationship helps healthcare providers detect minor shifts in a patient's behaviour or mood.

Self-Perceived Mental Health State

This result is assessed by asking participants the question:“In general, would you say your mental health is...?” using response options like “Poor,” “Fair,” “Good,” and “Excellent.” We coded the responses as binary variables, where the category “Poor Mental Health” encompasses responses of Poor and Fair and “Good Mental Health” includes responses of Good and Excellent. Self-perceived mental health is a widely recognized single-item measure that has demonstrated strong predictive validity for both current mental health status and long-term outcomes in emerging adults (Galambos et al., 2022). This measure shows strong reproducibility and sensitivity, which makes it a reliable indicator of overall well-being in large epidemiological studies.

Additional Data Collection

The database has a number of additional variables that give a detailed profile of the study population. The CCHS dataset includes demographic, socioeconomic, and health-related variables that allow for a comprehensive description of the emerging adult sample. We gathered data on variables including sex and age group to highlight important demographic differences. Sex is noted as a nominal variable (male or female), and age is split into two categories (18–19 years and 20–24 years) to investigate possible developmental differences during emerging adulthood. Additionally, we recorded the province of residence using standardized codes, which enabled us to explore the differences in healthcare access and outcomes across regions. We measured socioeconomic status by looking at household income to capture variations in economic background. We rated their level of education by asking participants whether they were currently enrolled in school, college, CEGEP, or university. Additionally, work status was recorded as a binary variable to show whether participants had been employed in the previous 12 months. The additional variables were crucial not only for outlining the sample but also for monitoring potential confounding factors in our statistical models. Prior research has consistently demonstrated that socioeconomic factors, such as household income and education level, significantly influence both access to primary healthcare and mental health outcomes (Adler & Ostrove, 1999; Marmot, 2005)

Statistical Analysis

Statistical analyses were conducted using the IBM Statistical Package for Social Sciences (SPSS). Data in Table 1 were organized by summarizing the demographic characteristics of the study population using descriptive statistics. Categorical variables were reported as counts and percentages (n(%)). To address the research question of whether access to a primary healthcare provider is associated with self-reported mental health status, a chi-square test with Yates' Correction for Continuity was performed to examine differences between groups. Mental health status was re-coded as either "poor mental health" or "good mental health", and access to a primary healthcare provider was re-coded as "yes" or "no". This re-coding allowed for a clear comparison of self-perceived mental health outcomes based on whether individuals had access to a primary healthcare provider. To assess the strength of the association observed in the chi-square test, effect size was evaluated using the Phi coefficient, calculated by the SPSS software. This measure provided additional context beyond statistical significance by indicating the magnitude of the association for a 2x2 chi-square contingency table. A p-value of less than 0.05 was set as the threshold for significance for the chi-square test. Effect sizes were interpreted using the guidelines set out by Akoglu (2018): Values around 0.05 were considered weak, 0.10 moderate, 0.15 strong, and 0.25 or greater indicated very strong associations (Akoglu, 2018).

Results

The database originally contained information collected from 109,470 respondents. After applying the inclusion criteria and filtering missing cases, the number of participants included in this study was reduced to 6819. Following participant filtering, the dataset was complete with no missing data, ensuring that all variables were accounted for in the final analysis. Table 1 shows the descriptive characteristics of the study population, the proportion of participants belonging to each demographic and primary variable. 807 participants self-reported poor mental health and 6012 reported good mental health. 75.7% of those with poor self-perceived mental health have a regular healthcare provider, and 70.1% of those with good self-perceived mental health have a regular healthcare provider. The results from a chi-square test with Yates Correction for Continuity are visualized in Figure 1. This test indicated a significant difference $\chi^2 = 10.737$, $p < 0.001$ in regular health care provider access between those who self-reported poor or good mental health. The calculated effect size using Phi's coefficient was determined to be $\Phi = 0.040$.

Table 1.

Descriptive Characteristics of survey participants (N=6819)

Characteristics	**Perceived Mental Health: Poor (n=807)**	**Perceived Mental Health: Good (n=6012)**
Sex		
Male	300 (37.1)	3071 (51.0)
Female	507 (62.8)	2941 (48.9)
Age		
Age between 18 and 19	191 (23.6)	1647 (27.4)
Age between 20 and 24	616 (76.3)	4365 (72.6)
Province		
Maritimes	124 (15.3)	576 (9.58)
Ontario	249 (31.0)	1707 (28.4)
Quebec	114 (14.2)	1383 (23.0)
Prairies	174 (21.4)	1398 (23.2)

British Columbia	113 (14.0)	751 (12.5)
Northern Territories	33 (4.08)	200 (3.32)
Total Household Income		
No Income or > $20 000	154 (19.0)	768 (12.8)
$20 000 - $39 999	138 (17.1)	903 (15.0)
$40 000 - $59 999	101 (12.5)	736 (12.2)
$60 000 - $79 999	71 (8.79)	676 (11.2)
$80 000 or more	335 (41.5)	2905 (48.3)
Current Enrollment in School/College/CEGEP/University		
Yes	338 (41.8)	2707 (45.0)
No	466 (57.7)	3250 (54.0)
Employment Status		
Yes	682 (84.5)	5200 (86.5)
No	122 (15.1)	769 (12.8)
Has a Regular Health Care Provider		

Yes	611 (75.7)	4216 (70.1)
No	196(24.2)	1796 (29.8)

Descriptive characteristics of study participants (N = 6819). Demographic information such as age, sex, province of respondent, household income, and work/study status was included. Characteristics regarding our primary outcomes were also reported, such as the status of regular health care provider. Data is presented as frequency (n(%)).

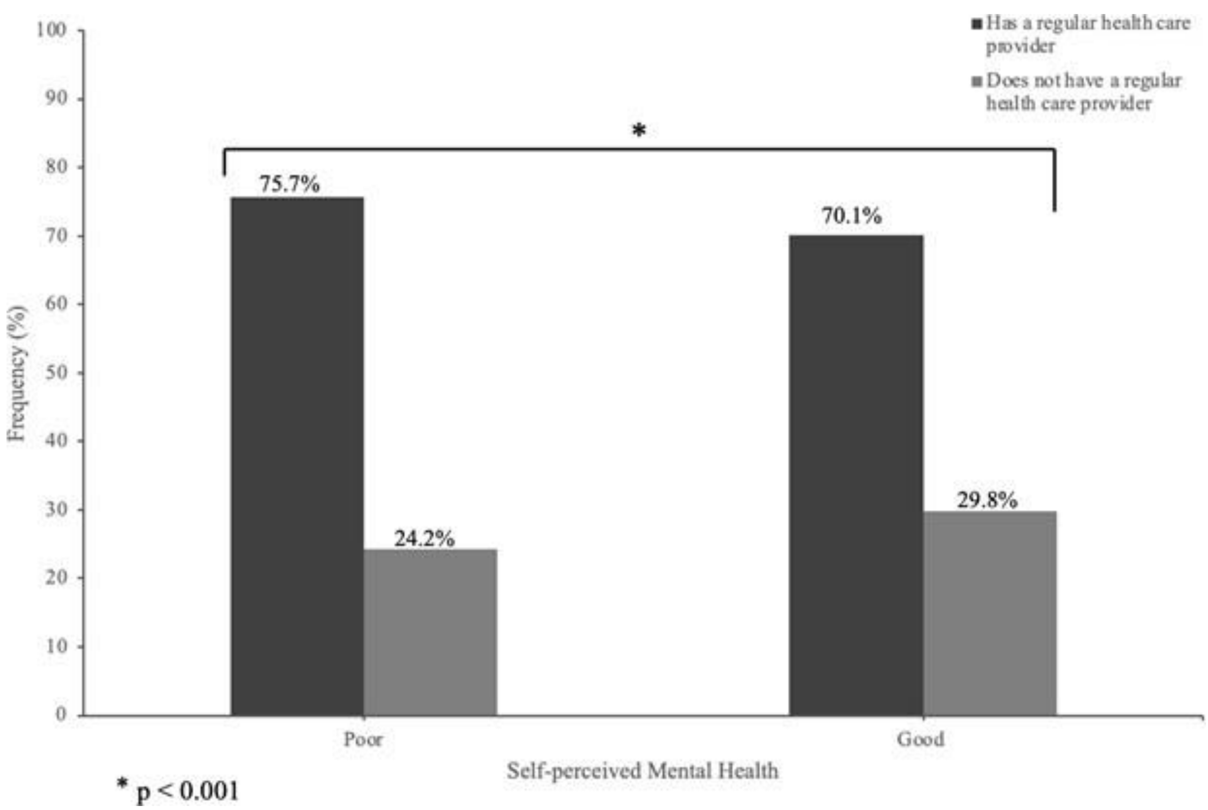


Figure 1. Frequency (%) of self-perceived mental health among those aged 18-24 in Canada with and without a regular health care provider. Data used in this chi-square test was collected from the Canadian Community Health Survey, a self-reported survey released in Canada every two years, where 6819 cases were used in this analysis. Evaluation of the frequency (%) of those who reported good or poor mental health and their access to a regular healthcare provider was obtained from the survey based on their answer to the questions “how is your mental health ?” and “Do you have a regular health care provider?”. A chi-square test with Yates’ Correction for Continuity was conducted in SPSS, where a p-value of < 0.001 was recorded.

Discussion

This study aimed to examine the association between access to a regular healthcare provider and self-perceived mental health among emerging adults (18-24 years old) in Canada. We found a statistically significant relationship between these two variables, showing that access to a regular healthcare provider is related to how people perceive their mental health; however, the effect size was weak. Interestingly, a slightly higher proportion of people with poor mental health had a regular provider, compared to those with good mental health.

Previous studies have consistently demonstrated that individuals without regular primary care are more likely to have an unmet mental health care need and poor mental health conditions (Awe et al., 2018; Miller-Matero et al., 2018; Oza et al., 2025). People with poor mental health often face a variety of barriers that hinder their access to primary healthcare, such as stigma, lack of mental health knowledge, and financial constraints, preventing emerging adults from seeking general care for mental health (Murray & Knudson, 2023). These barriers are commonly associated with reduced access to regular care among people with mental health challenges. However, results from our present study show a contrasting finding. We found that a greater proportion of individuals who self-reported poor mental health had access to a regular primary healthcare provider compared to those who reported good mental health. This result could indicate that there may be increased health-seeking behaviour among individuals with self-perceived mental health challenges. It is also a possibility that access to a regular healthcare provider does not always imply a higher quality or adequacy of mental health support received. For example, a previous study in 2022 showed that when patients received mental healthcare from a primary care provider, they rated their quality of care much lower than compared to mental healthcare received from a mental health specialist (Kyanko et al., 2022).

Several limitations need to be considered in our study. First, the subjective nature of self-reported mental health outcomes introduces potential response bias, since self-perceived mental health can be influenced by individual differences, health literacy, or contact with a healthcare provider. It may be the case that those with a regular provider or higher health literacy could be more aware of their mental health status and more likely to report as “fair” or “poor”, while those without access or with lower health literacy might not recognize their symptoms or underreport an accurate representation of their mental health status. This is shown in a study by Degan et al. (2019) in which they found that those with low health literacy had poorer mental health and higher psychological distress. This conclusion suggests that those with a higher health literacy may be better equipped to recognize and advocate for mental health needs. Second, our study only evaluated one primary independent variable without adjusting for other factors. This limitation may explain the finding of a small effect size; while access to healthcare can facilitate treatment and support of mental health, numerous other factors influence mental health, such as socioeconomic status, social supports, lifestyle, and pre-existing mental health conditions. Another limitation that may have influenced our study is the questions on the survey that we used to assess healthcare provider access. The variable we used only assessed whether or not they had a regular healthcare provider, but did not evaluate the frequency or quality of these regular healthcare visits. In addition, key limitations exist in our study due to the reliance on categorical data to conduct our statistical analysis. Recoding mental health status into binary variables for the chi-square test allowed us to test for an association between the two groups, but doing so may have oversimplified the data, preventing us from capturing the full range of mental health experiences.

Even with these limitations, our research presents a few new contributions. By concentrating on emerging adults in Canada, our findings offer insights into mental health trends: A greater percentage of emerging adults who view their mental health negatively have a regular primary healthcare provider, in contrast to those who view their mental health positively. This implies that people with worse mental health tend to get directed to primary care services more often or seek mental health treatments on their own. This trend underscores the significance of primary healthcare for detecting early onset of mental health symptoms and providing a time frame for clinical interventions and management. Our findings indicate that although a single-answer measure is useful, particularly in broad surveys, it should be complemented with further measures for a more thorough understanding. This is especially important considering the findings by Arias-de et al. (2020), in which they demonstrated how self-reported depression on its own might not be sufficient for clinical diagnosis. Our findings highlight the significant and positive outcomes of incorporating mental health services into primary health care and reinforce the importance of having integrated clinical services in policy implications (Flood et al., 2023). The implementation of accompanying self-rated questions with detailed clinical measures can help us to further understand the complexity of mental health status and the factors contributing to the negative perception of individuals' mental health state (Arias-de et al., 2020). In our study, the differences in population characteristics and the specific mental health measurements led to variations in the sensitivity and specificity metrics. This focus on demographics provides a deeper insight into how age is represented in mental health. As more emerging adults seek more comprehensive care in the primary health care settings, researchers can investigate how the timing of primary health care at early stages of mood disorders affects their views and reporting of their mental health. This unique timeframe brings in a temporal and situational aspect that enhances longitudinal studies. Additionally, better access to mental health in primary care settings offers a chance to improve continuity of care and support for this vulnerable group. Overall, our findings add new perspectives on the effectiveness of self-report measures. Public health campaigns should be focused on improving mental health literacy and decreasing stigma, especially in various primary health care settings. Targeted educational campaigns for normalizing specialized mental health services in primary healthcare can enhance self-management of mental health, which would support larger system-level improvements (Brown & Browne, 2024).

Future studies in this avenue should explore the quality of mental health care given within primary care settings. This would elucidate the nuances we found in our current study and understand if there are still unmet mental health needs even with a regular health care provider. In addition to exploring additional factors influencing mental health like socioeconomic status and lifestyle, future research could also incorporate measures that assess mental health literacy, stigma, or willingness to report symptoms, which could mediate the relationship between access and self-perceived mental health. Other future directions include using more detailed classifications of mental health or continuous variables to better understand the nuances of the relationship between mental health and primary healthcare access in emerging adults.

Conducting longitudinal studies in this topic can further explore how access to regular healthcare influences trends in mental health in emerging adults over time.

References

Adler, N. E., & Ostrove, J. M. (1999). Socioeconomic status and health: What we know and what we don't. *Annals of the New York Academy of Sciences*, *896*(1), 3–15.

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, *18*(3), 91–93. https://doi.org/10.1016/j.tjem.2018.08.001

Arias-de, J., Vilagut, G., Antoni Serrano-Blanco, Martín, V., Molina, A. J., Valderas, J. M., & Alonso, J. (2020). Accuracy of Self-Reported Items for the Screening of Depression in the General Population. International Journal of Environmental Research and Public Health, 17(21), 7955–7955. https://doi.org/10.3390/ijerph17217955

Awe, O. A., Udoka Okpalauwaekwe, Lawal, A. K., Ilesanmi, M. M., Feng, C., & Farag, M. (2018). Association between patient attachment to a regular doctor and self‐perceived unmet health care needs in Canada: A population‐based analysis of the 2013 to 2014 Canadian community health surveys. The International Journal of Health Planning and Management, 34(1), 309–323. https://doi.org/10.1002/hpm.2632

Brown, B., & Browne, D. T. (2024). Youth mental health in a Canadian community sample during COVID‐19: Exploring the role of perceived sense of belonging. Journal of Community Psychology, 52(6), 720–738. https://doi.org/10.1002/jcop.23125

Degan, T. J., Kelly, P. J., Robinson, L. D., & Deane, F. P. (2019). Health literacy in substance use disorder treatment: A latent profile analysis. *Journal of Substance Abuse Treatment*, *96*, 46–52. https://doi.org/10.1016/j.jsat.2018.10.009

Flood, C. M., Thomas, B., & McGibbon, E. (2023). Canada's primary care crisis: Federal government response. *Healthcare Management Forum*, *36*(5), 327–332. https://doi.org/10.1177/08404704231183863

Galambos, N. L., Johnson, M. D., & Krahn, H. J. (2022). Self-rated mental health in the transition to adulthood predicts depressive symptoms in midlife. Current Psychology, 42(34), 30223–30234. https://doi.org/10.1007/s12144-022-04081-z

Kyanko, K. A., A. Curry, L., E. Keene, D., Sutherland, R., Naik, K., & Busch, S. H. (2022). Does Primary Care Fill the Gap in Access to Specialty Mental Health Care? A Mixed Methods Study. *Journal of General Internal Medicine*, *37*(7), 1641–1647. https://doi.org/10.1007/s11606-021-07260-z

Marmot, M. (2005). Social determinants of health inequalities. *The Lancet*, *365*(9464), 1099–1104.

Matud, M., Díaz, A., Bethencourt, J., & Ibáñez, I. (2020). Stress and Psychological Distress in Emerging Adulthood: A Gender Analysis. Journal of Clinical Medicine, 9(9), 2859. https://doi.org/10.3390/jcm9092859

McAlpine, D. D., McCreedy, E., & Alang, S. (2018). The Meaning and Predictive Value of Self-rated Mental Health among Persons with a Mental Health Problem. Journal of Health and Social Behavior, 59(2), 200–214. https://doi.org/10.1177/0022146518755485

Miller-Matero, L. R., Khan, S., Thiem, R., DeHondt, T., Dubaybo, H., & Moore, D. (2019). Integrated primary care: patient perceptions and the role of mental health stigma. *Primary Health Care Research & Development*, *20*, e48. https://doi.org/10.1017/S1463423618000403

Murray, J. K., & Knudson, S. (2023). Mental health treatment and access for emerging adults in Canada: a systematic review. Frontiers in Public Health, 11. https://doi.org/10.3389/fpubh.2023.1088999

Oza, P., Monteau, R., Lu, W., & Monteau, R. (2025). Examining the association between access to a primary care provider and perceived unmet mental health treatment need among New York City adults. *Discover Public Health*, *22*(1), 26. https://doi.org/10.1186/s12982-025-00418-8

World Health Organization. (2018). *Integrating mental health into primary care: A global perspective*. https://www.who.int/mental_health/policy/primary_care/en

Grade Level, Sex, and Parental Smoking Status Significantly Influence Youth Smoking Rates in Canada: An Observational Study

Group 14 - Matthew Deng, Adrian Kalaikumar, Muhammad Noorani, Abhishek Sai Pendyala

McMaster University

School of Interdisciplinary Sciences

LIFE SCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8, 2025

Grade Level, Sex, and Parental Smoking Status Significantly Influence Youth Smoking Rates in Canada: An Observational Study

Introduction

Tobacco smoking is widely considered to be the leading cause of preventable and premature death worldwide, accounting for nearly 8 million deaths annually (Samet, 2013; WHO, n.d.). Much of the danger associated with cigarette smoking stems from the large amounts of toxic chemicals inhaled. The burning of cigarettes produces approximately 7,000 chemicals, 69 of which are carcinogenic (American Lung Association, 2024). The inhalation of nicotine, a highly addictive chemical compound found in tobacco leaves, worsens this issue. It stimulates the release of dopamine in the brain, leading to feelings of pleasure and the reinforcement of smoking behaviours (Centers for Disease Control and Prevention [CDC], n.d.). Abrupt cigarette cessation can trigger intense symptoms of withdrawal, highlighting the need for early interventions to prevent smoking initiation altogether (CDC, n.d.).

Adolescents, defined by the World Health Organization (WHO) as individuals aged 10-19, are disproportionately affected by the negative effects of cigarette smoking because their initiation occurs during a critical period of development (WHO, n.d.). The prefrontal cortex (PFC), located in the frontal lobe of the brain, is responsible for higher-order decision-making and does not fully mature in most cases until the age of 25 (Goriounova & Mansvelder, 2012). Nicotine can interfere with PFC maturation, leading to an increased risk of developing psychiatric disorders and impairments in impulse control, mood, and attention (CDC, n.d.). Furthermore, adolescents who initiate smoking are more likely to develop lifelong chronic addictions (American Lung Association, 2024). Of adults who smoke daily, 87% tried their first cigarette when they were 18 years old (American Lung Association, 2024). While smoking rates among Canadian youth aged 12-17 have declined from 2.5% in 2019 to 0.3% in 2022, disparities remain across key socio-demographic groups (Canadian Community Health Survey [CCHS], 2022). Gender-based differences in smoking prevalence have been well documented. In the early twentieth century, cigarette smoking in Western countries was predominantly seen as a male activity, resulting in only a small percentage of women smoking (Waldron, 1991). However, in more recent times, smoking rates between the sexes have converged, reflecting a broader societal acceptance of smoking among women (Waldron, 1991). Research by Jamal et al. (2016) showed that 16.7% of adult males in the United States engaged in cigarette smoking compared to 13.6% of females. These patterns may be rooted in physiological, behavioural, and social differences (Waldron, 1991). In addition to sex, another factor impacting youth smoking status is Indigenous status. The Indigenous population in Canada consists of the First Nations, Inuit, and Métis. Sikorski et al. (2019) reported that Indigenous youth in Canada had five times higher odds of becoming smokers compared to non-Indigenous youth. Similarly, in 2014-2015, the smoking rate among Aboriginal and Torres Strait Islander people aged 18 and over was 29% higher compared to the Non-Indigenous Australian population (Australian Bureau of Statistics, 2017). These statistics may be linked to the intergenerational trauma experienced by this group, with cigarette smoking functioning as a coping mechanism for managing these negative experiences (Elton-Marshall et al., 2011). Moreover, age plays a pivotal role in smoking initiation and progression. However, this study will extrapolate age from grade level. According to a report by the National Survey on Drug Use and Health (NSDUH, 2012), cigarette smoking among youth increased with age, from 1.2% in 12–13-year-olds to 13.6% in 16–17-year-olds. Likewise, research has shown that 29% of students in grades 10 to 12 have tried smoking cigarettes, compared to only 9% of students in grades 7 to 9 (Health Canada, 2024). Beyond socio-demographic characteristics, parental smoking status emerges as a notable predictor of adolescent smoking initiation (Gilman et al., 2009). Studies have shown that the presence of a parent who smoked significantly increased the probability of their child initiating smoking as well (Gilman et al., 2009). These findings may explain why some youth initiate smoking (Gilman et al., 2009).

Although data on the differences in smoking prevalence among socio-demographic groups such as sex, age, and Indigenous status have been documented, there is limited research specifically focused on adolescent populations. Most studies categorize adolescents as a single age group, hindering the potential to conduct intricate analyses of smoking trends across the adolescent developmental stages. Addressing this knowledge gap can aid public health officials and policymakers in developing grade-specific, tailored interventions geared toward lowering smoking rates among youth. The objective of this study will be to examine how socio-demographic factors, in particular, grade level, sex, and Indigenous status, are associated with youth smoking status among grade 6-12 students in public and private schools across all provinces in 2008-2009. This study will also touch on why some youth initiate smoking by exploring the relationship between parental smoking status and youth smoking status in this population.

Methods

This observational, retrospective, cross-sectional study utilized representative smoking behavior data from the 2008/2009 version of the Canadian Youth Smoking Survey (YSS), accessed via Odesi. The YSS is a cross-sectional national school-administered smoking survey conducted every two years intending to provide a baseline measurement for youth smoking and monitor the effectiveness of existing legislative policies on tobacco regulation (Elton-Marshall et al., 2011). A stratified, multistage random sampling design was implemented to ensure national representativeness (Elton-Marshall et al., 2011). Several schools were excluded from the sampling process, including those located in the Territories, specialized schools for students with special needs, schools on First Nation reserves, and those situated on military bases (Elton-Marshall et al., 2011). A total of 51,922 students completed the pen-and-paper survey, which was administered during class time and took approximately 30-40 minutes to finish (Elton-Marshall et al., 2011).

The study participants consisted of grade 6-12 students attending public and private schools across all 10 provinces in Canada from 2008-2009. All grade levels were included in the study to measure any changes in smoking prevalence as grade levels increased. No respondents were excluded from this study to maximize the sample size and ensure representative data.

Missing cases, indicated by values of 99 or “I do not know” for variables such as Indigenous status, parental smoking status, and the average number of cigarettes smoked on smoking days, were excluded from the analysis. This resulted in the population size decreasing to 43,719 respondents. Respondents identifying as First Nations, Inuit, or Métis for Indigenous status were grouped under a single variable called Indigenous. Similarly, for the variable smoking status, current and former smokers were combined to create a new value label called ever-smoked.

The primary outcome of this study was smoking status, where respondents could choose from either current smoker, former smoker, or never smoker. Smoking status was determined based on student responses to questions such as whether they had ever tried a cigarette if they had ever smoked a whole cigarette, if they had ever smoked 100 or more whole cigarettes in their lifetime, and how many of the last 30 days they had smoked one or more cigarettes (Elton-Marshall et al., 2011). Current smokers were defined as those who smoked 100 or more whole cigarettes and smoked at least once in the past 30 days (Elton-Marshall et al., 2011). Former smokers were defined as those who had smoked 100 or more whole cigarettes but had not smoked in the last 30 days (Elton-Marshall et al., 2011). Never smokers were defined as having never tried smoking in their life. Smoking status was selected as the primary outcome due to its high reproducibility, as it is defined based on a standardized set of principles and predetermined questions. For instance, the National Youth Tobacco Survey (NYTS), a national survey conducted in the United States to measure smoking behaviours among youth, uses similar guidelines for defining smoker status (CDC, 2024). In the NYTS, current smokers were defined as having smoked at least 100 cigarettes in their lifetime and within the past 30 days (CDC, 2024).

Two data sets were created; One containing only ever-smokers and the other containing never-smokers. Normality tests were conducted for the continuous variable, the average number of cigarettes smoked on smoking days for the ever-smoker population. The Shapiro-Wilk P value was less than 0.05, indicating that the data was not normally distributed. As a result, a Mann-Whitney U test was used to analyze the average number of cigarettes smoked on smoking days to sex, Indigenous status, and parental smoking status. A Kruskal-Wallis H Test was used to determine if there are statistically significant differences between grade level and province for the average number of cigarettes smoked on smoking days. Boxplots were created for the average number of cigarettes smoked on smoking days against all categorical variables. A 2x2 chi-square test was run to compare differences between parental smoking status and all categorical variables except for province for the current smoker population. Since the table was 2x2, Yate’s continuity correction value was used to assess significance. Bar graphs were made for the non-smoking population, showing the frequency of non-smokers for all categorical variables. The threshold for statistical significance was set at $p < 0.05$. All statistical analyses, box plots, and bar charts were created using IBM SPSS Statistics (Statistical Package for the Social Sciences), version 30. Table 1 was created using Microsoft Word.

Results

A total of N=43,719 cases were included in the dataset after data cleansing, with 3,141 ever-smokers and 40,001 never-smokers. The data were tested for normality and were found to be not normally distributed with a kurtosis>1. The following tables and figures were conducted using the procedures mentioned in the methods section. A total of 8,203 cases were omitted for having missing data. Full-sized charts are included in Appendix 1 for further reference.

VARIABLES		FREQUENCY (%)
GRADE	Grade 6	5813 (13.3)
	Grade 7	7536 (17.2)
	Grade	6807 (15.6)
	Grade 9	7199 (16.5)
	Grade 10	6838 (15.6)
	Grade 11	5648 (12.9)
	Grade 12	3878 (8.9)
PROVINCE	Prince Edward Island	3057 (7)
	Nova Scotia	3192 (7.3)
	New Brunswick	6453 (14.8)
	Quebec	6760 (15.5)
	Ontario	7939 (18.2)
	Manitoba	4427 (10.1)
	Saskatchewan	2466 (5.6)
	Alberta	2026 (4.6)
	British Columbia	3443 (7.9)
SEX	Male	21998 (50.3)
	Female	21727 (49.7)
INDIGENOUS STATUS	Indigenous	3718 (8.5)
	Non-indigenous	40001 (91.5)
SMOKING STATUS	Current Smoker	3125 (7.1)
	Former Smoker	16 (0.01)
	Never Smoker	40578 (92.8)
PARENTAL SMOKING	Parents smoke	18579 (42.5)
	Parents do not smoke	23878 (54.6)
	I do not know	1262 (2.9)
TOTAL	43719	

Note: Percentage of total proportion are show in parenthesis.

Table 1. The frequency of students in each variable and subcategory under grade, province, sex, Indigenous status, smoking status and parental smoking status (N=43719). It shows the counts and proportions of the number of cases in each category.

Table 1 highlights the number of students falling under each category for each variable. This data includes both ever-smokers and never-smokers, provides the count for the total cleansed dataset, and establishes a baseline for the study. The largest number of students is in grade 7, with 17.2% of the population, and the lowest is in grade 12, with 8.9%. The distribution of sex is nearly equal, comprising 50.3% males and 49.7% females. A disproportionate majority of students identified as non-Indigenous (91.5%). Moreover, most students' parents/step-parents/guardians are not smokers (54.6%). These descriptives provide the baseline values of the demographic data.

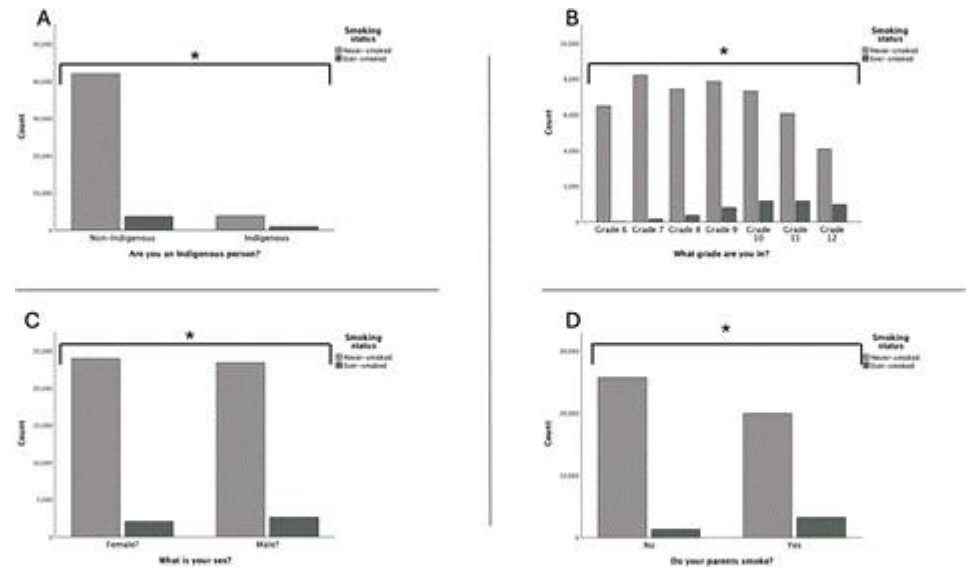


Figure 1. Clustered bar graphs displaying the count of individuals by smoking status across different demographic categories. Light grey bars represent never-smokers, and dark grey bars represent ever-smokers. An asterisk “*” denotes a statistically significant difference (p>0.05).

- Panel A: Smoking status by Indigenous status.
- Panel B: Smoking status by grade level (Grades 6–12).
- Panel C: Smoking status by sex.
- Panel D: Smoking status by parental smoking status.

Figure 1 visually compares the number of students who are ever-smokers and never-smokers across each variable of interest. Chi-square tests show a statistically significant association between smoking status and all demographic variables, with p-values less than 0.001.

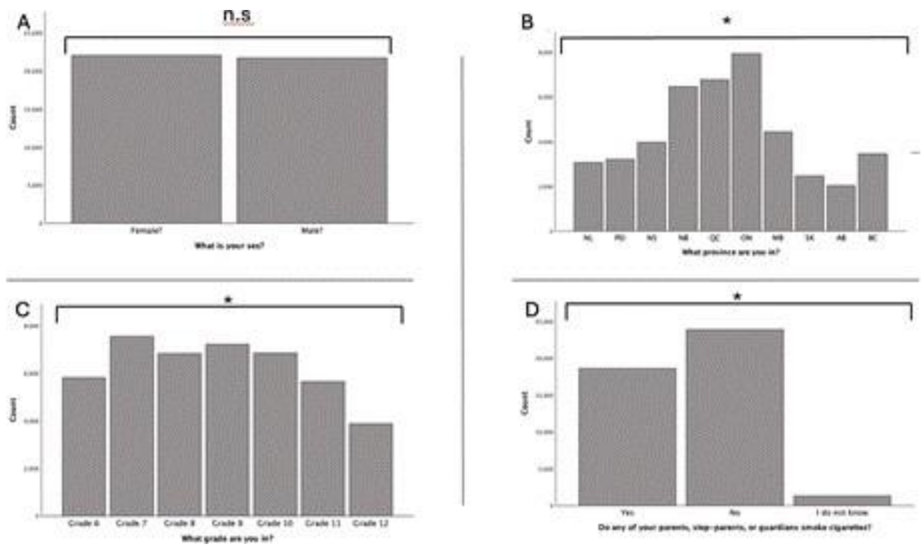


Figure 2. Simple bar graphs displaying the count of never-smokers for different demographic categories. An asterisk *"*"* denotes a statistically significant difference (p>0.05). An asterisk *"*"* denotes a statistically significant difference (p>0.05), *"n.s"* denotes a non statistically significant difference (p>0.05).

- Panel A: Count of never-smokers between males and females.Panel B: Count of never-smokers across provinces.
- Panel C: Count of never-smokers across grades (grades 6-12).
- Panel D: Count of never-smokers against parental smoking status.

Figure 2 represents the number of students who are non-smokers for each variable of interest. There is a near equal distribution in the frequency of never-smokers between the sexes, resulting in a non-statistically significant difference with a chi-square p-value <0.05. The number of smokers varies by region, with Ontario having the highest number of smokers, with 7,939 (18.2%) students and Alberta having the least with 2,026 (4.6%) students. Similar to the results in Table 1, grade 7 has the highest number of students and grade 12 has the least. Finally, most non-smokers' parents do not smoke either.

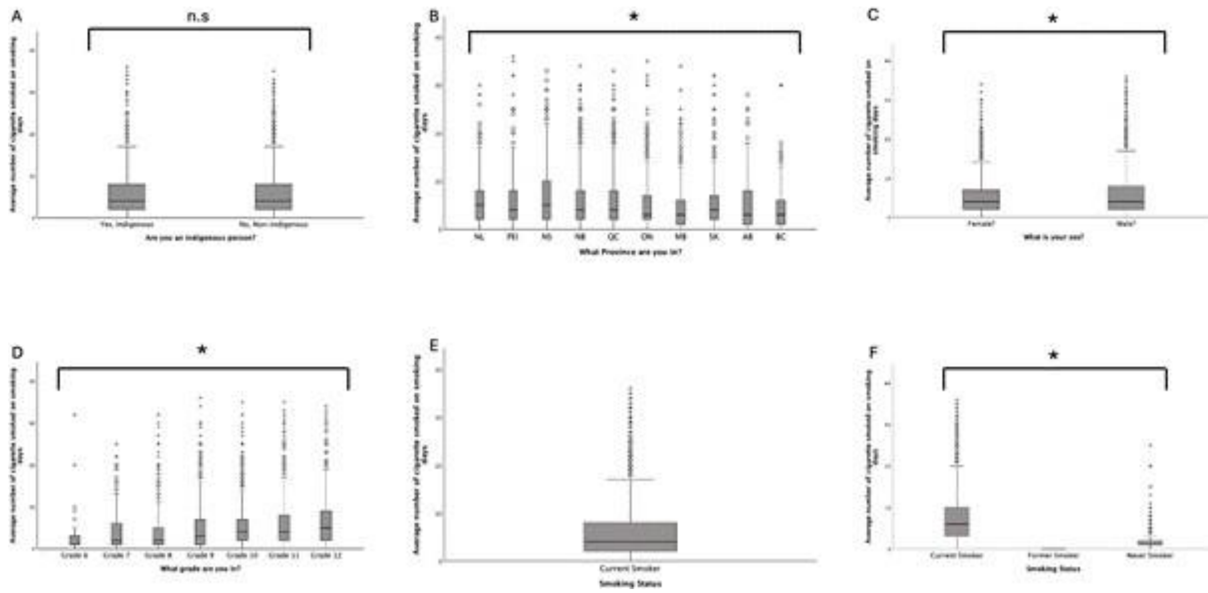


Figure 3. Simple box plots displaying and comparing the average number of cigarettes smoked on smoking days against various demographic factors for ever-smokers only. Circles represent moderate outliers (1.5<IQR<3.0), and triangles represent extreme outliers (IQR<3.0). An asterisk *"*"* denotes a statistically significant difference (p>0.05), *"n.s"* denotes a non statistically significant difference (p>0.05).

- Panel A: Smoking frequency by Indigenous status.
- Panel B: Smoking frequency by province (NL-Newfoundland and Labrador, PEI-Prince Edward Island, NS-Nova Scotia, NB-New Brunswick, QC-Quebec, ON-Ontario, MB-Manitoba, SK-Saskatchewan, AB-Alberta, BC-British Columbia).
- Panel C: Smoking frequency by sex.
- Panel D: Smoking frequency by grade level (grades 6-12).

- Panel E: Smoking frequency for current smoker.
- Panel F: Smoking frequency by smoking status.

Figure 3 represents the data points of ever-smoker students across each variable of interest. After conducting a Mann-Whitney U test for the dichotomous variables and a Kruskal-Wallace test for the multivariate variables, it was found that all of them except Indigenous status have a statistically significant difference between each category. Most variables had a significance value >0.001, showing a statistically significant difference, however, the significance value between Indigenous and non-Indigenous ever-smokers was 0.858, suggesting a high similarity of the median number of cigarettes smoked between these two groups.

PARENTAL SMOKING VERSUS	*ASYMPTOMATIC SIGNIFICANCE (2-SIDED)*
GRADE	<0.001
SEX	<0.001
INDIGENOUS STATUS	<0.001
TOTAL	4249
Note: Minimum significance value >0.05	

Table 2. The Chi-square asymptomatic significance value compares parental smoking status against grade, sex and Indigenous status (N=4249).

The results from Table 2 compare the strength of the difference in the expected values for grade, sex and Indigenous status for ever-smokers through the chi-square test. It shows that all variables have a statistically significant difference when grouped by parental smoking status.

Discussion

This study examined the association between socio-demographic factors, mainly grade level, sex, and Indigenous status and smoking status for Canadian youth in grades 6-12. The findings suggest that grade level, sex, and parental smoking status cause a significant difference in youth smoking prevalence. Indigenous status was not statistically significantly different from the average number of cigarettes smoked on smoking days for youth smokers. This is contrary to previous studies that have shown that youth smoking rates for Indigenous students are double that of non-Indigenous students (Elton-Marshall et al., 2011). This is an interesting difference, as both studies used the same dataset. However, the study by Elton-Marshall included students from grade 9 to 12 alone and compared it to 2,620 Indigenous and non-Indigenous students living off-reserve. This difference in the number of cases and different group methods may have caused this contradicting result. A study by Jetty (2017) found higher parental smoking levels, lower socioeconomic status and the cultural importance of tobacco to be the cause behind the higher smoking rates for Indigenous youth. The chi-square analysis between Indigenous status and parental smoking status shown in Table 2 suggests that parental smoking might have a bigger impact on Indigenous students than non-Indigenous students. Therefore, this study is partially in line with previous literature and reaches a valuable conclusion with regards to parental smoking status' impacts on Indigenous smoking.

Most of the findings of this study align with previous literature and strengthen the knowledge within this field. A study by the University of Waterloo (n.d) showed that there is a statistically significant difference in smoking rates between males and females in grades 6 to 9. The same conclusion has been reached in this study with a strong difference noted by the Mann-Whitney U test significance value of >0.001 in the smoking rates between male and female ever-smokers. This is a notable difference from the near equal split between the sexes in the never-smoker dataset and the combined dataset, further validating this disparity. This shows that the difference in sexes is specifically linked with youth smoking status and that males are more likely to smoke than females. Furthermore, a study by Gilman et al. (2009) showed that youth initiation of smoking differed by sex, with male kids being more likely to smoke given that their parents smoke. This result was particularly stronger with male children whose fathers smoked. This is in line with the findings of our study where the chi-square significance value <0.001 shows that there is a significant difference between parental smoking status and the youth's sex. Therefore, the results of this study can be interpreted to show that parental smoking is more likely to impact male youth than females. A study by Vuolo & Staff (2013) showed that parental smoking rate could increase the risk for youth smoking through poor support in education, psychological distress and a weakened attachment with the parents.

A study by the National Cancer Institute [NCI], (2020) showed that youth in High School (Grades 9-12) had a higher prevalence of cigarette smoking of 4.6% than youth in Middle School (Grades 7-8) with 1.6%. This is similar to the findings of this study, with a higher average of cigarettes smoked as grade increases, with students in Grade 12 having the highest median amount of cigarettes and Grade 6 having the least. The Kruskal-Wallace test compared the median smoking levels across grades and found the difference to be statistically significant with a significance value <0.001. This is further amplified by Figure 1, panel B, which shows that grade 12 students had the highest smoking rate despite making up the least proportion of the dataset, showing the increase in smoking rate as grade increases. Compared to the study by NCI, this study was able to provide more insight into the specific number of smokers by grade rather than a broad category.

This study has several limitations that must be acknowledged when interpreting the results. Firstly, a list-wise deletion method was used to remove all missing cases from the database. This corresponded to the loss of 8,203 responses, accounting for 15.8% of the total population. The remaining 84.2% of the study population may not have represented the actual target population, limiting the generalizability of findings beyond the sample. This method also introduced bias, where certain groups may be more prone to not answering one question, leading to their subsequent removal and underrepresentation. Future studies should use alternative methods for handling missing cases that do not reduce the sample size dramatically and minimize biases, such as multiple imputation, which predicts missing values (Austin et al., 2020). Secondly, the foundation of this study was inherently flawed by relying solely on self-reported data from the YSS. Self-reported data may be highly susceptible to recall bias, where respondents fail to remember details accurately. For example, the variable, average number of cigarettes smoked on smoking days, may have been harder to quantify if respondents smoked a large amount. To mitigate this inaccuracy, future studies could build off this study by implementing a longitudinal study design where respondents are given cigarette diaries to track daily smoking habits over a longer period. This may also assist in showcasing how smoking behaviours change over time and assess the effectiveness of specialized interventions for the same group of individuals.

This study provides valuable insights into the socio-demographic factors influencing smoking behavior among Canadian youth. The significant trends identified between sex, grade level, and parental smoking status demonstrate the need for targeted interventions in smoking prevention programs. Since sex plays a crucial role in smoking initiation, tailored strategies should provide a greater emphasis on sex-specific prevention methods. Additionally, the study emphasizes the importance of grade-specific interventions, as smoking initiation tends to increase with higher grade levels. Incorporating anti-smoking campaigns early on, when smoking rates start to rise, in particular, in grades 9 to 10, could help prevent the peaks in smoking rates at grades 10 to 12. Parental smoking status also emerged as a significant predictor of youth smoking, suggesting that parent-based interventions could be effective in reducing youth smoking rates. Public health initiatives should focus on parental education and smoking cessation programs to support families in creating smoke-free environments for their children. Overall, the findings suggest that a multi-faceted approach is key for developing effective smoking prevention strategies for reducing smoking prevalence among students in grades 6-12 across Canada.

References

American Lung Association. (2024, November 12). *Tobacco use among children and teens*. https://www.lung.org/quit-smoking/smoking-facts/impact-of-tobacco-use/tobacco-use-among-children

Austin, P. C., White, I. R., Lee, D. S., & van Buuren, S. (2021). Missing Data in Clinical Research: A Tutorial on Multiple Imputation. *Canadian Journal of Cardiology, 37*(9), 1322–1331. doi:10.1016/j.cjca.2020.11.010

Australian Bureau of Statistics. (n.d.). *Aboriginal and Torres Strait Islander peoples*. https://www.abs.gov.au/statistics/people/aboriginal-and-torres-strait-islander-peoples

Centers for Disease Control and Prevention. (n.d.). *Youth and tobacco use*. U.S. Department of Health & Human Services. https://www.cdc.gov/tobacco/php/data-statistics/youth-data-tobacco/index.html

Elton-Marshall, T., Leatherdale, S. T., Manske, S. R., Wong, K., Ahmed, R., & Burkhalter, R. (2011). Research methods of the Youth Smoking Survey (YSS). *Chronic diseases and injuries in Canada*, *32*(1), 47–54.

Gilman, S. E., Rende, R., Boergers, J., Abrams, D. B., Buka, S. L., Clark, M. A., Colby, S. M., Hitsman, B., Kazura, A. N., Lipsitt, L. P., Lloyd-Richardson, E. E., Rogers, M. L., Stanton, C. A., Stroud, L. R., & Niaura, R. S. (2009). Parental smoking and adolescent smoking initiation: an intergenerational perspective on tobacco control. *Pediatrics*, *123*(2), e274–e281. https://doi.org/10.1542/peds.2008-2251

Goriounova, N. A., & Mansvelder, H. D. (2012). Short- and long-term consequences of nicotine exposure during adolescence for prefrontal cortex neuronal network function. *Cold Spring Harbor perspectives in medicine*, *2*(12), a012120. https://doi.org/10.1101/cshperspect.a012120

Health Canada. (2020). *Summary of results for the Canadian Student Tobacco, Alcohol and Drugs Survey 2018-19*. Government of Canada. https://www.canada.ca/en/health-canada/services/canadian-student-tobacco-alcohol-drugs-survey/2018-2019-summary.htmlOpen Government+2Canada.ca+2Canada.ca+2

Jamal A, King BA, Neff LJ, Whitmill J, Babb SD, & Graffunder CM (2016). Current cigarette smoking among adults—United States, 2005–2015. MMWR Morbidity and Mortality Weekly Report, 65, 1205–1211.

Jetty R; Canadian Paediatric Society, First Nations, Inuit and Métis Health Committee, Ottawa, Ontario. Tobacco use and misuse among Indigenous children and youth in Canada. Paediatr Child Health. 2017 Oct;22(7):395-405. doi: 10.1093/pch/pxx124. Epub 2017 Sep 25. PMID: 29491724; PMCID: PMC5823001.

National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General. Atlanta (GA): Centers for Disease Control and Prevention (US); 2014. 13, Patterns of Tobacco Use Among U.S. Youth, Young Adults, and Adults. Available from: https://www.ncbi.nlm.nih.gov/books/NBK294302/

National Institute of Health. (n.d.). Youth tobacco use. Cancer Trends Progress Report. https://progressreport.cancer.gov/prevention/tobacco/youth_smoking

Samet J. M. (2013). Tobacco smoking: the leading cause of preventable disease worldwide. *Thoracic surgery clinics, 23*(2), 103–112. https://doi.org/10.1016/j.thorsurg.2013.01.009

School of Public Health Sciences. (n.d). *Smoking prevalence among males and females.* Retreived from the website of the University of Waterloo: https://uwaterloo.ca/tobacco-use-canada/youth-tobacco-use/current-smoking/smoking-prevalence/smoking-prevalence-among-males-and-females#:~:text=Smoking%20prevalence%20was%20not%20significantly,%2D9%20in%202018%2D19.&text=Among%20youth%20aged%2015%2D19%2C%20smoking%20prevalence%20was%20significantly%20greater,be%20reported%20(Figure%207.5).

Sikorski, C., Leatherdale, S., & Cooke, M. (2019). Tobacco, alcohol and marijuana use among Indigenous youth attending off-reserve schools in Canada: cross-sectional results from the Canadian Student Tobacco, Alcohol and Drugs Survey. Consommation de tabac, d’alcool et de cannabis chez les jeunes autochtones qui fréquentent des écoles hors réserve au Canada : résultats transversaux de l’Enquête canadienne sur le tabac, l’alcool et les drogues chez les élèves. *Health promotion and chronic disease prevention in Canada : research, policy and practice, 39*(6-7), 207–215. https://doi.org/10.24095/hpcdp.39.6/7.01

Vuolo, M., & Staff, J. (2013). Parent and child cigarette use: a longitudinal, multigenerational study. Pediatrics, 132(3), e568–e577. https://doi.org/10.1542/peds.2013-0067

Waldron I. (1991). Patterns and causes of gender differences in smoking. *Social science & medicine (1982)*, *32*(9), 989–1005. https://doi.org/10.1016/0277-9536(91)90157-8

World Health Organization: WHO. (2023, July 31). *Tobacco.* https://www.who.int/news-room/fact-sheets/detail/tobacco#:~:text=The%20tobacco%20epidemic%20is%20one,%2Dhand%20smoke%20(4)

The Association Between Cigarette Smoking and Difficulty Falling Asleep Within the Past Two Weeks Among Canadian High School Students (Grades 9-12) in the 2018-2019 School Year: ***A Comparative Analysis of Smokers and Non-Smokers***

Tori Edwards | Eunice Kim | Eunice Lee

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8, 2025

Introduction

Sleep is a fundamental pillar of health, influencing overall well-being and specifically playing a role in the development of adolescents (McCurry et al., 2024; Bonilla et al., 2024). However, sleep disturbances–including insufficient sleep duration, unusual sleeping behaviours, and daytime sleepiness–are becoming increasingly common in adolescents (McCurry et al., 2024). Globally, an estimated 7-36% of adolescents experience sleep disturbances, highlighting the growing concern surrounding adolescent sleep health (Liu et al., 2023).

Based on the existing research, various lifestyle factors–such as diet, physical activity, stress levels, and substance use–have been associated with poor sleep quality (Shochat, 2012). Among these factors, cigarette smoking has emerged as a potential contributor to sleep disturbances. Sun and Li (2024) analyzed data from the National Health and Nutrition Examination Survey (NHANES) in the United States and found that increased cigarette use over the past 30 days was correlated with a higher prevalence of diagnosed sleep problems, insufficient sleep duration, frequent snoring/snorting, sleep apnea, and daytime fatigue in smokers. Similarly, Yolten et al. (2010) examined sleep patterns in children aged 6–12 who were exposed to second-hand smoke at home and diagnosed with asthma. Their findings indicated that second-hand smoke exposure was associated with longer sleep onset latency, increased daytime sleepiness, and breathing irregularities during sleep. Other studies have reported similar associations between smoking and sleep difficulties. Nunez et al. (2020) found that individuals who smoked at night were 2.5 times more likely to experience moderate to severe insomnia than non-smokers. The same study found that smokers are 3.3 times more likely to experience very short sleep duration compared to non-smokers. Kang and Bae (2021) examined the relationship between cigarette and e-cigarette use in South Korea and found that dual users reported statistically significantly worse sleep quality than non-smokers. Their study also identified a correlation between smoking, sleep disturbances, and depression.

The impact of substance use on adolescent health has been a growing concern in Canada, particularly regarding smoking behaviours and their long-term health consequences. The New Brunswick Student Wellness Survey (NBSWS) provides insight into adolescent substance use, including smoking habits, and their impact on health outcomes (Government of New Brunswick, 2020). Conducted as part of the New Brunswick Health and Education Research Group (NBHERG), this survey collects data from students in grades 6–12 to assess health behaviours such as substance use, physical activity, and mental well-being (Government of New Brunswick, 2020). According to the 2018-2019 survey, 7% of students in grades 9–12 reported smoking cigarettes in the past 30 days, and 36% reported experiencing difficulty sleeping at least some of the time (Government of New Brunswick, 2020). While these findings suggest a possible link between smoking and sleep disturbances, existing research has yet to fully explore the relationship between cigarette smoking and sleep onset difficulties in Canadian adolescents.

Although prior studies have demonstrated an association between smoking and sleep disturbances, there remains limited research on how cigarette smoking specifically relates to difficulty falling asleep among adolescents. Additionally, no recent studies have investigated this relationship in Canadian high school students in grades 9-12. Much of the existing literature has focused on broader sleep outcomes, different age groups, or secondary exposures, such as second-hand smoke, rather than directly examining the link between active cigarette smoking and sleep onset difficulties. Given the increasing prevalence of sleep disturbances and smoking among adolescents, further research is necessary to clarify this relationship and its implications for adolescent health.

This study aims to examine whether self-reported cigarette smoking is associated with difficulty falling asleep within the past two weeks among Canadian high school students in grades 9–12 during the 2018-2019 school year. It is hypothesized that cigarette smokers in this age group will report a greater frequency of experiencing difficulty falling asleep compared to their non-smoking peers. Unlike previous studies that have explored multiple aspects of sleep quality, this research specifically investigated the association between cigarette smoking and sleep onset difficulty, an area that remains understudied.

Methods

This study is an observational, cross-sectional study using a database taken from Odesi, which allowed us to analyze a large sample at a specific time point. This database examined the Canadian Student Tobacco, Alcohol, and Drugs Survey (CSTADS) 2018-2019. The CSTADS collects information regarding tobacco use, drug and alcohol use, bullying, and sleep (Government of Canada). The CSTADS is a school-based survey designed to be nationally representative of grade 7-12 students. In New Brunswick (NB), the Propel Center works with the New Brunswick Health Center (NBHC) to run this survey concurrently with the New Brunswick Student Wellness survey (Department of Health & Public Health New Brunswick, 2021). Within each eligible school in NB, questionnaires were distributed among students using an 8:1 ratio of distribution, with the majority of students receiving the NBSWS and the remaining the CSTADS questionnaire (Burkhalter et al., 2019).

In 2018/2019, the survey ran from October 2018 to June 2019. The target population was young Canadians attending private, public, or catholic schools in grades 7-12, with a total of 116 school boards and 442 schools participating (Burkhalter et al., 2019). Groupings were categorized based on health region and type of school; within each province, two to three health region categories and two school-level categories were defined. Within each stratum, in each province, schools were selected based on simple random sampling (University of Waterloo, 2019). Two elementary schools for every secondary school were sampled to account for lower enrollments in elementary schools, ensuring appropriate distribution of participants. This stratified random sampling design created a representative population and minimized selection bias. Students who attended special schools (e.g., visually-impaired schools, hearing-impaired schools, daycares, special needs, First Nation reserve schools, virtual schools, schools located on military bases, international schools) were excluded from the target population to eliminate non-comparable data (Burkhalter et al., 2019). Participants who did not state their smoking status or difficulty falling asleep in the past two weeks were excluded from the study, as this data was needed for the analysis. Only non-smokers or smokers were included in the sample to isolate the effects of current smoking habits on sleep. We chose to remove former smokers because previous research has found that active smokers are two times more likely to have sleep disturbances than former smokers (Patterson et al., 2017).

Our study examined the results from high school students (grades 9-12) from ten different provinces: Newfoundland and Labrador, Prince Edward Island, Nova Scotia, New Brunswick, Québec, Ontario, Manitoba, Saskatchewan, Alberta, and British Columbia. All students in eligible grades at selected schools were allowed to participate. Students who attended special schools and students from schools that did not have at least 20 students enrolled in at least one eligible grade were excluded from the target schools (Burkhalter et al., 2019).

In each school, groups of students completed the CSTADS questionnaire and participated in focus groups immediately after. This data was self-reported with 65 questions in the original survey, and eight were used in our study. This included six to describe our population and two to answer the research question. The questionnaire was completed in a 40-minute window and was completely confidential. Several variables had an “I do not know” option, which was treated as “Not stated (99)”. This missing data was removed since we needed all responses from the participants to be included in the study. The independent variable was recorded as smoking status and was asked using three options: non-smoker, former smoker, or current smoker. For our study, this was collapsed down to non-smokers and smokers. A non-smoker is defined as someone who reports that they have not smoked 100 or more cigarettes in their lifetime but may have smoked a whole cigarette. A smoker is someone who has smoked at least 100 cigarettes in their lifetime and has smoked at least one cigarette in the past 30 days (Burkhalter et al., 2019). The primary dependent variable was difficulty falling asleep and was examined using the question: “In the last two weeks, how often have you had an extremely hard time falling asleep?”. Results were examined using the scale: 1: Every day/ night, 2: Several times, 3: Twice, 4: Once, 5: Never, 99: Not Stated (Burkhalter et al., 2019). This scale was then collapsed to be: 1: every day/night, 2-4:frequently, and 5: never.

We collected additional data to use as descriptive variables for our Table 1. This included participants' demographics such as grade, sex, province, and whether the school was located in an urban or rural area. We also examined the median household annual income of the area where their school was located, measured in Canadian dollars ($). Original responses ranged from $40,000 to $120,000, with responses being rounded to the nearest $5,000. This was then collapsed to three categories: under $60,000, $60,000-$80,000, and above $80,000. Finally, we decided to include a sleep descriptor variable to examine the sleep satisfaction of participants in the past two weeks. This was asked using the question, ‘In the last two weeks, how often have you felt satisfied with your sleep?' Responses were recorded on the same scale as our primary variable of ability to fall asleep.

To compare the groups in Table 1, we ran chi-squared analyses of our categorical data in SPSS to see if there were differences between the demographics of our groups. Associations between smoking and difficulty falling asleep were examined using a non-parametric Mann-Whitney U test due to the ordinal nature of our sleeping variable. This was run in SPSS, comparing the medians of each group with a significance threshold of $p<0.05$.

Results

The goal of this study was to examine the association between smoking status and self-reported difficulty falling asleep in Canadian high school students in Grades 9-12 in the 2018-2019 school year (n=35402). Table 1 shows that this study included participants across 10 provinces in Canada who were grouped in 2 categories: smokers (n=1662) and non-smokers (n=33740). A chi-squared test showed a significant difference in the proportions of smokers and non-smokers for grade, sex, province, school location, median household income, and sleep satisfaction ($p<0.001$). There was the highest number of smokers in grade 11 with 31% (517/1662) and the least number of smokers in grade 9 with 15% (253/1662). There were also more male smokers, with a percentage of 57% (950/1662), compared to female smokers, who had a percentage of 43% (712/1662). Across the 10 provinces, Quebec had the highest number of smokers, 17% (281/1662), while Manitoba had the least, 4% (62/1662). Smokers were also primarily living in an urban area, with a percentage of 63% (1038/1662) compared to a rural area, with 37% (624/1662). For median household annual income, the most number of smokers had a household income of $60,000-$80,000 with a percentage of 48% (804/1662), while a median household income of above $80,000 had the least number of smokers with a percentage of 15% (250/1662).

There was no missing data in the primary analysis because they were removed from the dataset since it would not be possible to examine the association between smoking and difficulty falling asleep without responses for smoking status and difficulty falling asleep. The original sample size from the survey was n=62850, but after removing non-high school students (grades 7 and 8) and participants who did not state smoking status or difficulty falling asleep, the new total sample size was n=35402, which consisted of n=1662 of smokers and n=33740 of non-smokers. A total of 27448 cases were removed from the dataset. Additionally, from the sample population (n=35402), n=7 from the smokers category and n=160 from the non-smokers category did not report sleep satisfaction in the past two weeks, and they were removed. The sample size for sleep satisfaction was n=1665 for smokers and n=33580 for non-smokers.

Table 1: Descriptive Variables of Study Participants in the 2018-2019 School Year (N=35402)

		Smoking Status		
		Smoker n=1662	Non-Smoker n = 33740	p-value
		n (%)	n (%)	
Grade	Grade 9	253 (15)	10651 (31)	<0.001
	Grade 10	428 (26)	10007 (30)	

	Grade 11	517 (31)	8286 (25)	
	Grade 12	464 (28)	4796 (14)	
Sex at birth	Female	712 (43)	16745 (50)	<0.001
	Male	950 (57)	16995 (50)	
Province	Newfoundland and Labrador	202 (12)	2887 (9)	<0.001
	Prince Edward Island	171 (10)	2463 (7)	
	Nova Scotia	253 (15)	3222 (9)	
	New Brunswick	188 (11)	1990 (6)	
	Québec	281 (17)	9121 (27)	
	Ontario	89 (5)	2531 (7)	
	Manitoba	62 (4)	1797 (5)	
	Saskatchewan	152 (9)	1860 (5)	
	Alberta	138 (8)	3900 (12)	
	British Columbia	126 (8)	3969 (12)	
School location	Urban	1038 (63)	25761 (76)	<0.001

	Rural	624 (37)	7979 (24)	
Median household annual income of the area where the respondent's school is located	Under $60,000	608 (37)	11395 (34)	<0.001
	$60,000-$80,000	804 (48)	14365 (42)	
	Above $80,000	250 (15)	7989 (24)	
Sleep satisfaction in the last two weeks[a]	Every day/night	258 (15)	4571 (14)	<0.001
	Frequently	989 (60)	24629 (73)	
	Never	408 (25)	4380 (13)	

Table 1. The proportions were compared for each variable using the Chi-squared test. [a]Missing data are described in the results section.

Figure 1 displays the percentage of individuals in the smoker group (n=1662) and non-smoker group (n=33740) for three different frequencies for difficulty falling asleep in the past 2 weeks (Every night, frequently, never). A Mann-Whitney U Test showed a statistically significant difference between the self-reported frequency of difficulty falling asleep of smokers and non-smokers. Smokers had significantly lower sleep quality than non-smokers for each category of frequencies (U = 33566904; p<0.001). For difficulty falling asleep every night, there was a higher percentage of individuals in the smoker group, with a percentage of 26% (438/1662), compared to the non-smoker group, which had a percentage of 12% (4000/33740). For the category, frequently, the percentage of individuals in the smoker group was lower than the non-smoker group, with percentages 55% (918/1662) and 59% (19913/33740 individuals), respectively. Lastly, for the category, never, the percentage of individuals in the smoker group was 18% (306/1662), which was lower than the non-smoker group, with a percentage of 29% (9827/33740).

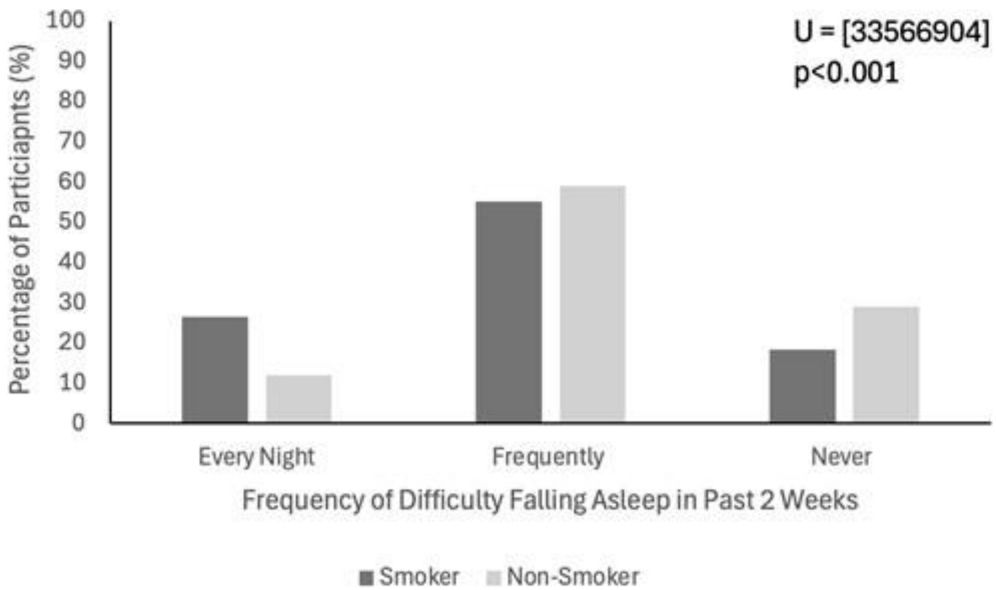


Figure 1. Frequency of Difficulty Falling Asleep for Smokers and Non-Smokers. This figure displays the percentage of individuals in the smoker group (n=1662) and non-smoker group (n=33740) for three different frequencies of difficulty falling asleep in the past 2 weeks: every night, frequently, and never. The total sample size was n=35402, and there was a significant difference between the medians of the smoker and non-smoker groups (U = 33566904; $p<0.001$).

Discussion

Our findings found that self-reported smoking is associated with the ability to fall asleep in the past two weeks in Canadian high school students. It was seen that smokers experienced a higher frequency of difficulty falling asleep than non-smokers. This supports the idea that smoking negatively affects sleep onset in high schoolers. These findings are consistent with previous studies, such as Sun and Li (2014), who found that smoking was correlated with various sleeping problems. Our findings are also consistent with the study by Yolten et al. (2010), who found that these smoke-related sleeping difficulties were also present in children rather than only adults. Interestingly, our study found that more non-smokers frequently had difficulty falling asleep compared to smokers. This may have been due to combining sleep categories, but more research would be needed to examine this inconsistency in our results.

The findings of this study have implications for both individual health and public health interventions. For adolescent smokers, understanding the negative impact of smoking on sleep quality may motivate them to adopt healthier habits. Given that high school students are still in a crucial stage of physical and cognitive development, prioritizing sleep health is essential for their overall well-being (Bonilla et al., 2024). Poor sleep quality has been linked to reduced academic performance, increased mental health challenges, and a higher risk of developing chronic health conditions later in life (Sun & Li, 2024).

From a public health perspective, these findings highlight the need for targeted interventions aimed at reducing smoking among adolescents. School-based education programs that emphasize the connection between smoking and sleep disturbances could be an effective strategy to discourage smoking initiation. Additionally, healthcare providers working with adolescents should incorporate sleep assessments into routine check-ups and guide them on improving sleep hygiene for students who report difficulties falling asleep.

Overall, while this study presents valuable insights into the relationship between smoking and sleep quality, future research should aim to address its limitations through objective sleep assessments, broader sampling strategies, and consideration of additional influencing factors. The implications of this research underscore the importance of prioritizing sleep health in adolescent populations and integrating these findings into public health initiatives aimed at reducing smoking and promoting overall well-being.

While this study's findings provide valuable insights into the relationship between smoking and sleep quality among Canadian high school students, some limitations must be acknowledged. First, the study relies on retrospective self-reported data collected over the past two weeks. This method introduces potential bias, as participants may not accurately remember or report their sleep quality. Self-reported measures are vulnerable to social desirability bias, where participants may either overestimate or underestimate their sleep quality to align with perceived norms (Nikolopoulou, 2022).

Another key limitation is that sleep quality was assessed based on participants' perceptions rather than objective measures. Subjective assessments are prone to individual differences in interpretation, meaning that two students with similar sleep patterns may report different sleep quality ratings based on personal experiences and expectations. Some participants may have provided false responses, either intentionally or unintentionally, which could skew the results.

Data limitations also arise due to missing responses regarding smoking status, leading to the exclusion of certain participants from the analysis. This exclusion reduces the overall sample size and may introduce selection bias, as students who did not disclose their smoking status could differ in important ways from those who did. Additionally, the study sample does not encompass all Canadian high school students, limiting the generalizability of the findings. According to provincial education statistics, Ontario alone has approximately 660,856 high school students, yet only around 2000–3000 students from the province participated in the survey

(*Facts about Elementary and Secondary Education*, n.d.). This discrepancy highlights the limited representativeness of the study sample.

Another notable limitation is the exclusion of students from Canada's three territories: Yukon, Northwest Territories, and Nunavut. Due to unique socio-cultural and environmental factors, high school students in these regions may have different sleep patterns and smoking behaviours. Without their inclusion, the findings may not fully reflect the experiences of all Canadian high school students.

Finally, this study does not account for other factors that could influence sleep quality beyond smoking behaviour. Variables such as academic stress, family dynamics, and physical activity levels all play a role in sleep health and could confound the observed relationship between smoking and sleep disturbances. Future studies should aim to control for these additional influences to obtain a more comprehensive understanding of the factors affecting adolescent sleep.

Different study designs could be considered to address these limitations and improve the accuracy of future research. One potential approach is to conduct a sleep study utilizing objective measures such as electroencephalography (EEG) or actigraphy (Cudney et al., 2021). These methods would allow researchers to collect precise data on sleep duration and sleep efficiency, minimizing the biases associated with self-reported sleep quality.

Additionally, future research could expand the sample size and ensure a more representative dataset by including students from a wider geographic range, particularly those in underrepresented regions such as Canada's northern territories. A longitudinal study design would also be beneficial in establishing causal relationships between smoking and sleep disturbances by tracking changes in sleep quality over time.

Finally, future studies should consider incorporating additional variables such as academic stress, socioeconomic status, and physical activity to better understand adolescent sleep health. By addressing these gaps, future research can provide a more comprehensive picture of the interplay between smoking and sleep in Canadian high school students.

References

Bonilla, D. A., Kočí, J., Petro, J. L., & Kreider, R. B. (2024). The Healthy Trinity to Face Non-communicable Diseases: Physical Activity, Nutrition and Sleep. *Integrated Science*, 89–121. https://doi.org/10.1007/978-3-031-64292-0_5

Burkhalter, R., Thompson-Haile, A., Cooke, M. (2019). 2018/2019 Canadian Student Tobacco, Alcohol and Drugs Survey Microdata User Guide. Waterloo, Ontario: Propel Centre for Population Health Impact, University of Waterloo, 1-39.

Cudney, L. E., Frey, B. N., McCabe, R. E., & Green, S. M. (2021). Investigating the relationship between objective measures of sleep and self-report sleep quality in healthy adults: a review. *Journal of Clinical Sleep Medicine*, *18*(3). https://doi.org/10.5664/jcsm.9708

Department of Health & Public Health New Brunswick. (2021). *Youth and Cannabis in NB: A Public Health perspective*. https://www2.gnb.ca/content/dam/gnb/Departments/h-s/pdf/en/Addiction/Youth-and-Cannabis-in-NB_PH-Perspective-2018-20.pdf

Facts about elementary and secondary education. (n.d.). Ontario.ca. https://www.ontario.ca/page/facts-about-elementary-and-secondary-education

Jung, J., & Fenelon, A. (2024). Do later school start times improve adolescents' sleep and substance use? A quasi-experimental study. *Preventive Medicine*, *185*, 108028–108028. https://doi.org/10.1016/j.ypmed.2024.108028

Kang, S.-G., & Bae, S.-M. (2021). The Effect of Cigarette Use and Dual-Use on Depression and Sleep Quality. *Substance Use & Misuse*, *56*(12), 1869–1873. https://doi.org/10.1080/10826084.2021.1958855

Liu, W., Wang, D., Fan, Y., Ye, H., Bu, L., & Fan, F. (2023). Changes in sleep disturbance among Chinese adolescents: A population-based online survey. *Journal of Psychosomatic Research*, 175, 111515. https://doi.org/10.1016/j.jpsychores.2023.111515

McCurry, K. L., Toda-Thorne, K., Taxali, A., Angstadt, M., Hardi, F. A., Heitzeg, M. M., & Sripada, C. (2024). Data-driven, generalizable prediction of adolescent sleep disturbances in the multisite Adolescent Brain Cognitive Development Study. Sleep, 47(6), zsae048. https://doi.org/10.1093/sleep/zsae048

Nikolopoulou, K. (2022, June 24). *What Is Social Desirability Bias? | Definition & Examples.* Scribbr. https://www.scribbr.co.uk/bias-in-research/what-is-social-desirability-bias/

Nunez, A., Rhee, J. U., Haynes, P., Chakravorty, S., Patterson, F., Killgore, W. D. S., Gallagher, R. A., Hale, L., Branas, C., Carrazco, N., Alfonso-Miller, P., Gehrels, J.-A., & Grandner, M. A. (2020). Smoke at night and sleep worse? The associations between cigarette smoking with insomnia severity and sleep duration. *Sleep Health*, *7*(2). https://doi.org/10.1016/j.sleh.2020.10.006

Patterson, F., Grandner, M. A., Malone, S. K., Rizzo, A., Davey, A., & Edwards, D. G. (2017). Sleep as a target for optimized response to smoking cessation treatment. *Nicotine & Tobacco Research*, *21*(2), 139–148. https://doi.org/10.1093/ntr/ntx236

Shochat, T. (2012). Impact of lifestyle and technology developments on sleep. *Nature and Science of Sleep*, 19. https://doi.org/10.2147/nss.s18891

Sun, H., & Li, S. (2024). Exploring the relationship between smoking and poor sleep quality: a cross-sectional study using NHANES. *Frontiers in Psychiatry*, *15*. https://doi.org/10.3389/fpsyt.2024.1407741

University of Waterloo. (2019). *About CSTADS/YYS*. Tobacco Use in Canada. https://uwaterloo.ca/tobacco-use-canada/about/data-sources/about-cstadsyss#:~:text=CSTADS%2FYSS%20provides%20timely%20and,data%20through%20classroom%2Dbased%20surveys.

Yolton, K., Xu, Y., Khoury, J., Succop, P., Lanphear, B., Beebe, D. W., & Owens, J. (2010). Associations Between Secondhand Smoke Exposure and Sleep Patterns in Children. *PEDIATRICS*, *125*(2), e261–e268. https://doi.org/10.1542/peds.2009-0690

-

Investigating the Association Between Work-Related Stress and Depression Among Working-Age Canadians (18–64 years): Comparing Those With and Without Depression and Exploring Illicit Drug Use as a Coping Mechanism

Group 16: Arshia Ojha, Julian Chan, Miriam Samuel

School of Interdisciplinary Science, Mcmaster University

LIFESCI 3LX3: Applied Biostatistics and Scientific Writing

Dr. Janet Pritchard

April 8th, 2025

Introduction

Depression is a common mood disorder characterized by increased irritability, decreased energy levels, reduced concentration, and mental changes that can severely reduce one's overall functionality (Chand & Arif, 2023). This disorder affects 280 million people globally and can cause drastic personality changes (World Health Organization, 2023). Depression is responsible for both individual suffering and interpersonal strain; repercussions of depression impact quality of life, including relationships, ability to look after oneself, and work performance/productivity (Leka & Jain, 2010). Additionally, depression, and other mental health disorders, place a huge strain on the Canadian economy, an estimated $50 billion per year (*Making the Case for Investing in Mental Health in Canada*, n.d.). This economic loss is attributed to mental health-related absences and healthcare expenses, with depression as a leading cause. In relation, around 13%-15% of Canadians experienced at least one major depressive episode in their lifetime as of 2022 (Stephenson, 2023), marking it a pressing issue of concern.

Occupational stress, or work stress, has become a rising cause of concern across the globe; there is an increasing number of employees who are being subjected to longer working hours, and more physically, and mentally demanding work-related tasks (Chireh et al., 2023). Additionally, 47% of working Canadians considered work to be the most stressful part of their day. Occupational stress can be attributed to the accumulation of demanding work-related tasks beyond an employee's capability to cope, and lack of access to proper emotional support resources, and is found to be a significant risk factor for depression by Madsen et al. (2017).

A vast collection of literature has identified a connection between occupational stress and depression: specifically occupational stress as a risk factor for depression (Habtu et al., 2024; Jung et al., 2023). Despite an array of current literature, this connection has not been thoroughly examined in Canada's diverse working populace and usually focuses on a narrow age range or specific occupations underlining a critical knowledge gap. For instance, a study conducted by Habtu et al. found occupational factors to be a responsible party in higher rates of depression, however, their specific finding applies to Ethiopian healthcare workers (2024). 60% of participants fell in the 27-30 age range, so as a result, their research was unevenly focused. Focusing on such a narrow range creates very specific research that is not generally applicable and excludes the vast majority of the working population. Jung et al. also found a positive correlation between occupational stress and symptoms of depression, however, they sampled participants from 18 government organizations and private companies, which limits the generalizability of their results, instead of focusing on broad occupational application (2023). Given Canada's reported stressful work environments (Sethumadhavan & Sasangohar, 2020), Canadian workers are substantially at risk for developing depression.

Coping mechanisms to manage occupational stress, specifically illicit drug use, is seldom explored with the intent of examining their correlation. Although de Fatima Fernandes and da Silva Gherardi-Donato discuss substance abuse as a coping mechanism, their focus is mainly on alcohol abuse (2017). Additionally, they discuss drug abuse generally, however considering the prevalence of illicit drug use in the Canadian population aged 12 and older, it is important to examine illicit drug use in the working population: 9.8 million Canadians, had used illicit drugs, 3.1 million of which had used within the past year (Bragazzi et al., 2021) Not all drug users abuse drugs, so it is also important to examine drug use in general instead of drug abuse specifically, as this will produce more generalizable results. Despite the prevalence of illicit drug use in the Canadian population, research prioritizes licit drugs as coping mechanisms and fails to address the specific risks of illicit drugs, such as abuse, overdose, and death in severe cases.

Research Questions

Our research addresses the outlined knowledge gaps by using the 2019-2020 Canadian Community Health Survey (CCHS) to investigate two research questions using the PECO (population, exposure, comparison, outcome) format:

1. **Primary Research question:** In working Canadians aged 18-64, how is occupational stress associated with depression?

2. **Secondary Research Question**: In working Canadians aged 18-64, is work stress correlated with illicit drug use in the past 12 months as a coping mechanism?

By answering these research questions, our goal is to successfully inform public health policies that create healthier, more cohesive workplaces, not only improving mental health but also reducing stigma surrounding mental health disorders.

Methods

Study Design

This study used the Canadian Community Health Survey (CCHS) - Annual Component from 2019-2020. The database is a file produced by the CCHS combining two years of data (Abacus Data Network, 2024). The CCHS is a cross-sectional survey that collects information related to health status, healthcare use, and health determinants for Canadians across the country (Statistics Canada, 2024). The objective of this survey is to support health surveillance programs by providing health data as a single data source for health research. To address emerging issues related to the health of the population in Canada, this survey creates a flexible instrument that facilitates a rapid response to address any issues. After a redesign implemented in 2015, the

CCHS now uses two different frames: an area frame and the Canadian Child Tax Benefit (CCTB) frame. Using the area frame, a sample of dwellings is selected to target the population aged 18 and over. The CCTB frame is used to sample persons aged 12 to 17 years. One child is then pre-selected to complete the survey. The size of the sample is enlarged during the selection process to account for non-responses and units outside the coverage (for example, vacant dwellings, institutions, children not eligible due to age or death, etc.) (Government of Canada, 2019).

Study Participants

The original database had 108,252 cases including missing cases. The inclusion criteria for our sample were: (1) Males and Females, (2) Adults aged 18-64 years in the workforce, (3) Individuals who answered the 'work stress', 'depression', and 'illicit drug use' questions in the CCHS. Our exclusion criteria were: (1) Respondents younger than 18 years or older than 65 years, (2) Missing cases or respondents who answered questions with 'prefer not to say', and (3) cases from variables not of interest in this study. A total number of 1921 cases was used for the analysis after missing values and inapplicable responses were excluded. To ensure we had data to analyze, not every missing case for each variable was excluded (see Table 1).

Study Outcomes

Our primary research question investigates the association between work stress and depression. The independent variable is work stress. The data collected was categorical as respondents answered the perceived work stress question with either 'Not at all stressful', 'A bit stressful', 'Not very stressful', 'Quite a bit stressful', or 'Extremely stressful'. We then grouped these categories into Yes Work Stress or No Work Stress. The No Work Stress group included those who answered with 'Not at all stressful' and the Yes Work Stress group included those who responded with 'A bit stressful', 'Not very stressful', 'Quite a bit stressful', and 'Extremely stressful' The dependent variable is depression. Respondents answered 'Minimal depression', 'Moderate depression', 'Mild depression', 'No depression', or 'Severe depression' to the depression severity question, which classifies the collected data as categorical. We grouped these categories into Yes Depression and No Depression. The Yes Depression group included 'Minimal depression', 'Moderate depression', 'Mild depression', and 'Severe depression', whereas the No Depression group included 'No depression'. This outcome was chosen to address the current knowledge gap in research, which is characterized by a lack of research investigating work-related stress and depression in Canadian working adults. This outcome is important because there is literature supporting work stress as a precipitator of depression – a previous study that measured work stress as job strain and low coworker support discovered that over a two-year period, men and women with high work stress had elevated odds of incident depression (Shields, 2006).

Our secondary research question investigates the association between work stress and illicit drug use. While the independent variable is still work stress, the dependent variable is illicit drug use. The data collected was categorical as respondents answered whether or not they have taken any illicit drugs in the past 12 months with either 'Has used at least one of the drugs' or 'Has not used any of the drugs'. These two categories were later renamed to Yes Illicit Drug Use and No Illicit Drug Use respectively. This outcome was chosen based on evidence highlighting that illicit drugs may be more likely to be used in response to work stress (Frone, 2008). This information is vital in promoting healthier coping strategies and supportive work environments in the future.

Additional Data Collection

Demographic data was collected from respondents including sex at birth, age, highest level of education, perceived life stress, perceived mental health, and satisfaction with life in general. These data have been included to characterize the population sample to provide vital information to help with comparing findings across studies. This type of data can also be used to determine whether certain demographic groups are disproportionately associated with certain occurrences (Hughes et al., 2016); this is important because demographic data can be used to make data-driven social, political, and economic conclusions (Call et al., 2023).

Statistical Analysis

There are no continuous data in this study. The chi-squared test was used to assess for differences between the Yes Work Stress and No Work Stress groups. This test was performed using the 29th version of the Statistical Package for Social Sciences (SPSS). Yate's Correction for Continuity was used as there was a comparison of categorical data between two groups in a two-by-two table. The Chi-squared test was chosen to determine if a difference between observed data and expected data is due to chance or if it is due to a relationship between work stress and either depression or illicit drug use. Demographic data are presented in Table 1 and data comparing outcomes between groups are presented in Table 2. A p-value of <0.05 was considered significant.

Results

To investigate the association between depression and work stress in this study's primary research question, results were analyzed on SPSS statistics using a chi-square analysis and plotted into bar graphs. 108,502 individuals were surveyed, between men and women in Canada with the age range between 18-64. However, due to missing data, only 1921 individuals' responses were considered valid and used. Displayed below is Table 1 which includes the descriptive characteristics of the database that were asked of participants.

Table 1

Baseline characteristics of study participants, stratified by work stress (N=1921).

Descriptive Variable	**Yes Work Stress (N=1765)**	**No Work Stress (N=156)**	**p-value (Chi-squared test)**
Sex at birth - no. (%)			
Male	828 (46.92%)	84 (53.85%)	0.114
Age - no. (%)*			
18 to 34 years	628 (35.58%)	57 (36.54%)	
35 to 49 years	589 (33.37%)	39 (25.00%)	0.061
50 to 64 years	548 (31.06%)	60 (38.46%)	
Highest level of education - no. (%)*			
Less than secondary school	35 (1.98%)	5 (3.21%)	
Secondary school only	384 (21.76%)	21 (13.46%)	0.430
Post-secondary certificate/ diploma/ university degree	1323 (74.96%)	119 (76.28%)	
Satisfaction with life in general - no. (%)*			

Very Satisfied	651 (36.89%)	80 (51.28%)	
Satisfied	996 (56.44%)	65 (41.67%)	0.007**
Neither satisfied nor dissatisfied	64 (3.63%)	6 (3.85%)	
Dissatisfied	44 (2.49%)	3 (1.92%)	
Very Dissatisfied	6 (0.34%)	1 (0.64%)	
Perceived mental health - no. (%)*			
Poor	38 (2.15%)	1 (0.64%)	
Fair	132 (7.48%)	6 (3.85%)	
Good	498 (28.22%)	40 (25.64%)	0.009**
Very good	672 (38.08%)	53 (33.97%)	
Excellent	421 (23.85%)	56 (35.90%)	
Perceived life stress - no. (%)*			
Not at all stressful	93 (5.27%)	56 (35.90%)	
Not very stressful	384 (21.76%)	40 (25.64%)	
A bit stressful	853 (48.34%)	44 (28.21%)	<0.001**
Quite a bit stressful	371 (21.02%)	12 (7.69%)	
Extremely stressful	62 (3.51%)	4 (2.56%)	
Grouped Income - no. (%)			
< $20,000	126 (7.14%)	16 (10.26%)	
$20,000 to $59,999	368 (20.86%)	44 (28.21%)	0.021
$60,000+	1271 (72.02%)	96 (61.54%)	

Results were collected from the Canadian Community Health Survey - Annual Component (CCHS), 2019-2020. 108,252 individuals answered questions related to their health status, health care utilization, and health determinants for the Canadian population. Individuals ranging from 12 to 65+ years old answered questions on descriptive characteristics. Individuals who did not state answer for all questions were excluded from this table, decreasing our study population from 108,252 to 1921 except for 'Highest Level of Education', 'Perceived mental health', 'Satisfaction with life in general', and 'Perceived life stress', who actual frequencies are N=1787, 1917, 1916, and 1919 respectively. The data is expressed by the units of n (number of individuals) and percentages (%) of the total surveyed population. Significance at the 0.05 level is denoted by *.

Furthermore, after the primary question data was collected, the study's secondary research question database was analyzed to determine if there was an association between work stress and illicit drug use along with the coexistence of depression. Table 2 outlines the different questions asked among participants working Canadians aged 18-64, including if they are depressed, stressed from work, and if they have used illicit drugs in the past 12 months. Of the questions asked, 1765 responses were inputted, and analyzed with a 2x2 chi-square test.

Table 2

Prevalence of depression and illicit drug use in those who have and have not experienced work stress

Outcome Variable	Yes Work Stress (N=1765)	No Work Stress (N=156)	p-value (Chi-squared test)
Depression - no. (%)			
Yes Depression	1208 (68.44%)	87 (55.77%)	0.002**
No Depression	557 (31.56%)	69 (44.23%)	
Illicit drug use - no. (%)			

Has used at least one drug in the last 12 months	67 (3.80%)	7 (4.49%)	0.834
Has not used any drugs in the last 12 months	1696 (96.08%)	149 (95.51%)	

Results were collected from the Canadian Community Health Survey - Annual Component (CCHS), 2019-2020. 10,252 individuals answered questions related to their health status, health care utilization, and health determinants for the Canadian population. Individuals who did not state answers for all questions were excluded with 'Illicit drug use' as an exception, whose actual frequency is 1919 respectively. The data is expressed by the unit n (number of individuals) and percentages (%) of the surveyed population. Significance at the 0.05 level is denoted by **.

When the primary research question investigated the association between work stress and depression, 1295 individuals reported feeling depressed and of those 1295 individuals that were depressed, 1205 reported feeling stressed from work which is 93% of the depressed population. The chi-square analysis results reported a p-value of .002 with the degree of freedom equaling one, inferring that there is statistical significance of association between work stress and depression since it is lower than p = 0.005. Figure 1, shown below, illustrates a bar graph comparing the outcomes of work stress and depression in working Canadians from ages 18-64.

Figure 1

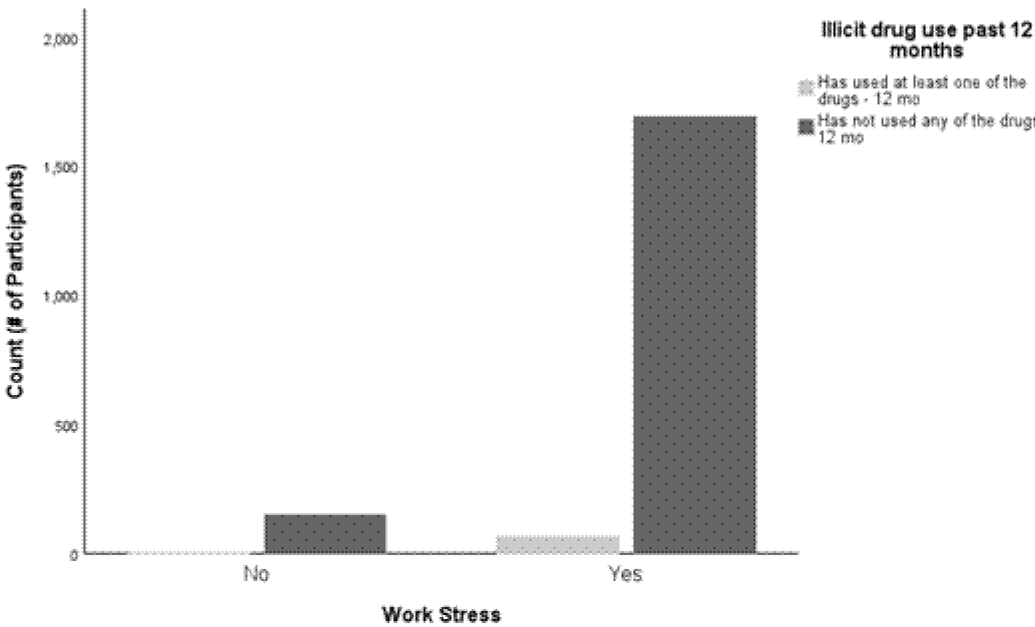
Outcomes of Work Stress and Depression in Working Canadians Aged 18-64

Responses were inputted to investigate the association between work stress and depression (N=1921). 1295 reported feeling depressed, with 1205 out of the 1295 reported being stressed from work.

In the second research question investigating work stress and illicit drug use, 1765 people reported feeling stressed in their work, out of which 67 reported taking illicit drug use in the past 12 months. 2x2 Chi-square results have yielded a P value of 0.834 with a degree of freedom also equalling one, which means there is no statistical significance because it is greater than P=0.005. Figure 2, displayed below, outlines the outcomes of illicit drug use within 12 months and work stress in working Canadians aged 18-64.

Figure 2

Outcomes of Illicit Drug Use Within 12 Months and Work Stress in Working Canadians Aged 18-64



Responses were inputted to investigate the association between work stress, and individuals. 1765 people reported feeling work stress, and 67 of the 1765 have reported taking illicit drugs while under work stress in the past 12 months.

Discussion

We found that there was a significant difference between those who experienced work stress and those who experienced depression. This reinforces the conclusion that work stress is associated with depression. Alternatively, we found that there was no statistical difference between those who experienced work stress and those who used at least one illicit drug in the past 12 months, leading us to the conclusion that work stress is not associated with illicit drug use. This absence of a statistically significant association between work stress and illicit drug use in this study can be explained by a few factors. Firstly, illicit drug use is a complex and multifaceted construct. It is possible that the measures CCHS used did not fully capture the variability of respondents' experiences for example, self-reported drug use may have been underreported due to stigma (Latkin et al., 2017). Moreover, illicit drug use is influenced by a range of intersecting factors such as mental health status, social support, and socioeconomic status (Sinha, 2008).

Compared to other published findings, our results are similar in that work stress is shown to be associated with depression (Letvak et al., 2012; Shields, 2006; Stansfeld et al., 1999). However, a limitation of these studies includes self-reported data which introduces potential false data and bias (Brenner & DeLamater, 2016). Alternatively, in terms of our secondary outcome, illicit drug use, our findings are different from the published literature that shows an association between work stress and illicit drug use (Frone, 2008). This difference may have arisen in light of the complex nature of work stressors. Work overload for example, as a work stressor is related to drug use because it partly contributes to an individual's self-evaluation of being a successful employee (Frone et al., 1994) and also extends into their self-image as a successful provider, spouse, or parent. If that self-image is threatened, they are more likely to turn to illicit drug use (Frone, 2008).

Given that published literature considers and highlights the nature of work stress, this may have led to the establishment of different results due to the psychosocial implications of work stress (Frone, 2008). Regardless of similarities and differences, this study contributes to a growing field of research investigating work stress and its relationship with mood disorders like depression and coping strategies like illicit drug use. Our data highlights that work stress is associated with depression but not with illicit drug use. Despite this, these findings can still inform future research directions and highlight the need for further exploration with alternative methodologies, larger samples, or different subpopulations.

Limitations

Some limitations of this study included the lack of job diversity in the questionnaire. The study does not account for the fact that different jobs may have different kinds of work stress as well as different levels. An example of this includes the difference in stress experienced between workforces such as manual labour or the healthcare workforce. Another limitation that was taken account for in this study, was the missing variables that were not accounted for. Out of 108,252 responses that were collected, only 1921 responses were valid data. There is a large number that is still unaccounted for, and this may skew our data in the study. Further studies may need to address this missing data gap to reduce it. Finally, a third limitation is that the study did not account for gender and ethnicity. Different genders and ethnicities may have different levels of depression and work stress. Another potential limitation is cultural responses. Different cultures respond to stress in different ways, as mediated by increased social interaction, for example.

Future Directions

Our findings highlight the need for continued research on the relation between occupational stress and depression, as well as coping mechanisms used to self-treat these conditions, in diverse populations. Future research should emphasize longitudinal studies to examine the temporal relationship between these variables. Cross-sectional data may not allow causal data interpretations. Monitoring stress levels and depression outcomes over time could clarify the order of these experiences: whether the sampled population experienced depression after experiencing occupational stress, or vice versa. This would help illuminate implementable preventative methods for employers to reduce negative mental-health-related outcomes. Utilizing a mixed-method design could also provide more thoughtful insights surrounding occupational stress and coping mechanisms, for example using qualitative interviews and quantitative surveys to create more thorough data and eliminate measurement imprecisions/inaccuracies. Qualitative interviews could provide insight into the reasons people choose drug-related coping mechanisms instead of healthier options and certain barriers people may face when trying to access mental health services. As the sample results of our data set were self-reported, future studies could benefit from medical health records for diagnosed conditions and utilizing biomarkers, such as cortisol to measure stress, to strengthen the significance of the data. Additionally, future studies could examine the role of emerging workplace trends and technologies in conjunction with occupational stress: including hybrid work models and the use of artificial intelligence in the workplace. Research exploring how these emergencies influence mental health outcomes could help with the formation of an adaptive workplace policy that supports mental health improvements and positive coping mechanisms.

Implications for Healthcare and Public Health

The findings highlight the importance of integrating mental health services into the workplace: including access to counseling, co-worker support networks, and techniques for managing stress. Workplace-healthcare partnerships could introduce and/or improve harm reduction protocols, such as providing naloxone training to assist in the case of an opioid overdose. Our findings can encourage employers to reform existing workplace policies to include days off work for mental health and mandatory workshops centered around engaging activities to improve mental health (these workshops could also inform about warning signs of occupational stress/depression and drug use/abuse in co-workers); these interventions can help to reduce risks of stress-related outcomes like depression, while simultaneously decreasing stigma surrounding mental health conditions and drug use. Decreasing stigma encourages people to seek out help for their conditions, promoting a healthier, more cohesive working environment. Public education on these topics should focus on accurate identification and timely interventions to prevent negative coping mechanisms from developing and encourage positive, inclusive attitudes towards those with mental health struggles. By confronting these gaps in our research, post-dating research and policy can aim to create strong, supportive working environments to reduce occupational stress, ultimately improving the burden of depression and substance use in Canadian society.

References

Abacus Data Network. (2024). *Canadian Community Health Survey—Annual Component (CCHS) 2019-2020* [Dataset]. Abacus Data Network. https://abacus.library.ubc.ca/dataset.xhtml?persistentId=hdl:11272.1/AB2/DMJPVY

Bhui, K., Dinos, S., Galant-Miecznikowska, M., de Jongh, B., & Stansfeld, S. (2016). Perceptions of work stress causes and effective interventions in employees working in public, private and non-governmental organisations: A qualitative study. *BJPsych Bulletin*, *40*(6), 318-325. doi.org/10.1192/pb.bp.115.050823

Bragazzi, N. L., Beamish, D., Kong, J. D., & Wu, J. (2021). Illicit drug use in Canada and implications for suicidal behaviors, and household food insecurity: Findings from a large, nationally representative survey. *International Journal of Environmental Research and Public Health*, *18*(12), 6425. doi.org/10.3390/ijerph18126425

Brenner, P. S., & DeLamater, J. (2016). Lies, damned lies, and survey self-reports? Identity as a cause of measurement bias. *Social Psychology Quarterly*, *79*(4), 333–354. doi.org/10.1177/0190272516628298

Call, C. C., Eckstrand, K. L., Kasparek, S. W., Boness, C. L., Blatt, L., Jamal-Orozco, N., Novacek, D. M., & Foti, D. (2023). An ethics and social justice approach to collecting and using demographic data for psychological researchers. *Perspectives on Psychological Science*, *18*(5), 979–995. doi.org/10.1177/17456916221137350

Canadian Centre for Occupational Health and Safety. (2016). *Mental health*. Canadian Centre for Occupational Health and Safety. Retrieved February 7, 2025, from https://www.ccohs.ca/topics/wellness/mentalhealth

Chand, S. P., & Arif, H. (2023). Depression. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK430847/

Chireh, B., Essien, S. K., Novik, N., & Ankrah, M. (2023). Long working hours, perceived work stress, and common mental health conditions among full-time Canadian working population: A national comparative study. *Journal of Affective Disorders Reports, 12*. doi.org/10.1016/j.jadr.2023.100508

de Fatima Fernandes, M. N., & da Silva Gherardi-Donato, E. C. (2017). Is it workplace stress a trigger for alcohol and drug abuse? *Open Journal of Nursing*, *7*(3).doi.org/10.4236/ojn.2017.73034

Frone, M. R. (2008). Are work stressors related to employee substance use? The importance of temporal context in assessments of alcohol and illicit drug use. *Journal of Applied Psychology*, *93*(1), 199–206. doi.org/10.1037/0021-9010.93.1.199

Frone, M. R., Cooper, M. L., & Russell, M. (1994). Stressful life events, gender, and substance use: An application of tobit regression. *Psychology of Addictive Behaviors*, *8*(2), 59–69. doi.org/10.1037/0893-164X.8.2.59

Government of Canada (2019, November 18). *Canadian Community Health Survey—Annual Component (CCHS)*. https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=1208978

Habtu, Y., Kumie, A., Selamu, M., Harada, H., & Girma, E. (2024). Prevalence and determinants of occupational depression, anxiety, and stress among Ethiopian healthcare workers. *Scientific Reports*, *14*(21817). doi.org/10.1038/s41598-024-72930-x

Hughes, J. L., Camden, A. A., Yangchen, T., & College, A. S. (2016). Rethinking and updating demographic questions: Guidance to improve descriptions of research samples. *Journal of Psychological Research*, *21*(3). https://www.rwu.edu/sites/default/files/demographic%20recs.pdf

Jung, S., Young, C. S., Lee, M. Y., Oh, K.-S., Shin, D.-W., Kim, E. S., Kim, M.-K., Jeon, S.-W., & Cho, S. J. (2023). Occupational stress and depression of Korean employees: Moderated mediation model of burnout and grit. *Journal of Affective Disorders, 339*, 127-135. doi.org/10.1016/j.jad.2023.07.045

Latkin, C. A., Edwards, C., Davey-Rothwell, M. A., & Tobin, K. E. (2017). The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. *Addictive Behaviors*, *73*, 133–136. doi.org/10.1016/j.addbeh.2017.05.005

Leka, S., & Jain, A. (2010). Health impact of psychosocial hazards at work: An overview. *World Health Organization*. https://affinityhealthhub.co.uk/d/attachments/health-impact-of-psychosocial-hazards-at-work-an-overview-1476975602.pdf

Letvak, S., Ruhm, C. J., & McCoy, T. (2012). Depression in hospital-employed nurses. *Clinical Nurse Specialist CNS*, *26*(3), 177–182. doi.org/10.1097/NUR.0b013e3182503ef0

Madsen, I. E. H., Nyberg, S. T., Magnusson Hanson, L. L., Ferrie, J. E., Ahola, K., Alfredsson, L., Batty, G. D., Bjorner, J. B., Borritz, M., Burr, H., Chastang, J.-F., de Graaf, R., Dragano, N., Hamer, M., Jokela, M., Knutsson, A., Koskenvuo, M., Koskinen, A., Leineweber, C., … Kivimäki, M. (2017). Job strain as a risk factor for clinical depression: Systematic review and meta-analysis with additional individual participant data. *Psychol Med*, *47*(8), 1342-1356. doi.org/10.1017/S003329171600355X

Making the Case for Investing in Mental Health in Canada. (n.d.). Mental Health Commission of Canada. Retrieved April 8, 2025, from https://mentalhealthcommission.ca/wp-content/uploads/2021/09/Investing_in_Mental_Health_FINAL_Version_ENG_0.pdf

Sethumadhavan, A., & Sasangohar, F. (2020). Design for health: Applications of human factors. *Elsevier Science*. doi.org/10.1016/C2018-0-00043-2

Shields, M. (2006). Stress and depression in the employed population. *Health Reports*, *17*, 11-29. https://pubmed.ncbi.nlm.nih.gov/17111591/

Sinha, R. (2008). Chronic stress, drug use, and vulnerability to addiction. *Annals of the New York Academy of Sciences*, *1141*(1), 105-130. doi.org/10.1196/annals.1441.030

Stansfeld, S. A., Fuhrer, R., Shipley, M. J., & Marmot, M. G. (1999). Work characteristics predict psychiatric disorder: Prospective results from the Whitehall II Study. *Occupational and Environmental Medicine*, *56*(5), 302-307. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1757742/

Statistics Canada. (2024). *Canadian Community Health Survey, 2019-2020: Annual Component*. Borealis. doi.org/10.5683/SP3/ZVCGBK

Stephenson, E. (2023). Mental disorders and access to mental health care. *Statistics Canada: Insights on Canadian Society*, *75-006-X*(2291-0840). https://www150.statcan.gc.ca/n1/pub/75-006-x/2023001/article/00011-eng.htm

World Health Organization. (2023, March 31). *Depressive disorder (depression)*. World Health Organization Fact Sheet. Retrieved February 7, 2025, from https://www.who.int/news-room/fact-sheets/detail/depression