medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Title**: The Individual and Social Determinants of COVID-19 Diagnosis in Ontario, Canada: A Population-Wide Study

**Authors**: Maria E. Sundaram, PhD[1], Andrew Calzavara, MS[1], Sharmistha Mishra*, MD, PhD[2,3,4,5], Rafal Kustra, PhD[6,7], Adrienne K. Chan, MD[2,3,6,8], Mackenzie A. Hamilton, BS[1,6], Mohamed Djebli, BS[1,6], Laura C. Rosella, PhD[1,6], Tristan Watson, MPH[1,6], Hong Chen, PhD[1,6,9], Branson Chen, MSc[1], Stefan D. Baral*, MD[10], Jeffrey C. Kwong*, MD[1,6,9,11,12]

Author affiliations:
[1] ICES, Toronto, ON
[2] Department of Medicine, University of Toronto, Toronto, ON
[3] Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON
[4] Institute of Medical Science, University of Toronto, Toronto, ON
[5] MAP Centre for Urban Health Solutions, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto
[6] Dalla Lana School of Public Health, University of Toronto, Toronto, ON
[7] Department of Statistical Sciences, University of Toronto, Toronto, ON
[8] Sunnybrook Health Sciences Centre, Toronto, ON
[9] Public Health Ontario, Toronto, ON
[10] Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MN
[11] Department of Family and Community Medicine, University of Toronto, Toronto, ON
[12] University Health Network, Toronto, ON
* Equal contribution

Corresponding author: Jeff Kwong, MD; jeff.kwong@utoronto.ca



**NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.**

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

## Abstract

**Background**: Optimizing the public health response to reduce coronavirus disease 2019 (COVID-19) burden necessitates characterizing population-level heterogeneity of COVID-19 risks. However, heterogeneity in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) testing may introduce biased estimates depending on analytic design.

**Methods**: We explored the potential for collider bias and characterized individual, environmental, and social determinants of testing and diagnosis using cross-sectional analyses among 14.7 million community-dwelling people in Ontario, Canada. Among those diagnosed, we used separate analytic designs to compare predictors of: 1) individuals testing positive versus negative; 2) symptomatic individuals only testing positive versus testing negative; and 3) individuals testing positive versus individuals not testing positive (i.e., testing negative or not being tested). Analyses included tests conducted between March 1 and June 20, 2020.

**Results**: Of a total of 14,695,579 individuals, 758,691 were tested for SARS-CoV-2, of whom 25,030 (3.3%) tested positive. The further the odds of testing from the null, the more variability observed in the odds of diagnosis across analytic design, particularly among individual factors. There was less variability in testing by social determinants across analytic designs. Residing in areas with highest household density (adjusted odds ratio [aOR]: 1.86; 95%CI: 1.75-1.98), highest proportion of essential workers (aOR: 1.58; 95%CI: 1.48-1.69), lowest educational attainment (aOR: 1.33; 95%CI: 1.26-

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

1.41), and highest proportion of recent immigrants (aOR: 1.10; 95%CI: 1.05-1.15) were consistently related to increased odds of SARS-CoV-2 diagnosis regardless of analytic design.

**Interpretation**: Where testing is limited, risk factors may be better estimated using population comparators rather than test-negative comparators. Optimizing COVID-19 responses necessitates investment and sufficient coverage of structural interventions tailored to heterogeneity in social determinants of risk, including household crowding, occupation, and structural racism.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

## Introduction

The spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus causing coronavirus disease 2019 (COVID-19), has resulted in a pandemic with heterogeneity in exposure and transmission risks.[1-4]

Heterogeneity in social determinants of COVID-19 may exist at the individual and network (for example, by housing density[5-7]) levels. In addition, social determinants of health, including barriers to healthcare, occupation, structural racism, and xenophobia have been implicated in COVID-19 risk.[8,9] Environmental determinants such as ambient air pollution may also play a role, as existing evidence indicates that higher ambient air pollution increases risk for infection with other respiratory viruses[10,11] and severe COVID-19.[12,13] It could also play a role in COVID-19 risk by operating within the context of low-quality housing and structural racism.[14,15]

Using observational data to identify determinants of COVID-19 relies on SARS-CoV-2 testing, a service which is not equally distributed.[16] Differential testing introduces the potential for selection biases,[17,18] including collider bias.[18] Collider bias may be introduced into epidemiological studies of COVID-19 risk factors if the factors under investigation are related to infection and to the likelihood of being tested.[18-20] For example, data suggest individuals with diabetes are more likely to develop severe COVID-19 if infected.[21,22] Thus, if infected, individuals with diabetes may be more likely to be tested, and consequently diabetes may appear to be associated with a diagnosis of COVID-19 in studies of individuals tested for SARS-CoV-2, even if diabetes is not a

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

mechanistic risk factor for infection.[18] The opposite may occur with underlying respiratory diseases (e.g., asthma) that have symptoms similar to those caused by SARS-CoV-2, leading to the appearance of potentially 'protective' associations with COVID-19.[23]

Our objectives were to: 1) explore the potential for collider bias in large-scale studies of COVID-19 determinants; and 2) examine individual, environmental, and social determinants associated with testing and diagnosis among 14.7 million individuals in Ontario, Canada.[18]

## Methods

### *Study population, setting, and design*

We conducted an observational study using population-based laboratory and health administrative databases from Ontario. Ontario's single-payer health system provides universal access to hospital and physician services[24] and laboratory testing.[25] We used data from individuals tested between March 1 and June 20, 2020 to identify determinants associated with testing, and then used different analytic designs to identify determinants associated with diagnosis.

### *Data sources, linkages, and inclusion criteria*

We identified testing status using data from the Ontario Laboratories Information System (OLIS), and linked this information to relevant health-related datasets containing

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

demographic, healthcare use, and area-level information. These datasets were linked using unique encoded identifiers and analyzed at ICES.[26]

OLIS captured approximately 88% of all laboratory-identified COVID-19 reported by the province during the study period (calculated by cases identified in OLIS divided by the number of cases reported by Ontario's COVID-19 dashboard in the same time frame). OLIS records included specimen collection date, results, and a text field for symptoms completed by healthcare providers at the time of sampling. We obtained individual- and area-level demographic and environmental information from the Registered Persons Database; the Canadian Institute for Health Information's Discharge Abstract Database, Same Day Surgery Database, and National Ambulatory Care Reporting System; the Ontario Health Insurance Program; the Ontario Mental Health Reporting System; the Ontario Population Health and Environment Cohort; and the 2016 Canadian Census[27] (**Supplemental Table 1**).

For individuals with more than one test in OLIS, we used the first positive or indeterminate test, or the first negative test if all tests during the study period were negative. Never-tested individuals during the study period were included if they were not recorded as deceased before, or born after, March 1, 2020. To assess determinants of testing and diagnosis, we included individuals who were tested for SARS-CoV-2 by polymerase chain reaction tests and were not residing in a long-term care facility as of March 1, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

*Selection and definition of potential determinants of COVID-19*

Individual-level determinants included sex, age group, underlying health conditions, and prior healthcare use. Underlying health conditions were selected as: 1) health conditions identified in the literature as associated with COVID-19 severity[2,28-31] or associated with symptoms similar to those caused by COVID-19, because severity and symptoms may lead to differential testing and thus, collider bias;[32-37] or 2) health conditions such as dementia that increase the need for personal care support, thereby reflecting an intersection with occupational risks among essential care providers.[38,39]

Healthcare use was hypothesized to increase access to testing, and/or signal a marker for comorbidities, and was measured by number of hospital admissions in the past 3 years, number of outpatient physician visits in the past year, and influenza vaccination in the 2019-2020 season. We also included ACG® System[40] Aggregated Diagnostic Groups (ADGs)[41] as a composite measure of comorbidities.

Environmental determinants included fine particulate matter ($PM_{2.5}$) using satellite-derived estimates[42] and land-use regression model for $NO_2$,[43] at the postal-code level.

Social determinants were conceptualized as area-based variables that may signal contact rates in communities [household density, apartment building density, uncoupled (e.g., not married) status];[44,45] contact rates at work ("essential workers");[17,46] socio-economic barriers to healthcare access and/or housing (household income and educational attainment);[47,48] and factors related to race/ethnicity (visible minority status

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

and recent immigration).[8,9] These variables were derived from the 2016 Canada Census at the level of dissemination areas (DA), the smallest geographic unit that Census data are disseminated.[49] DAs were ranked at the city level (for median per-person income equivalent) or at the province level (for all other social determinants), and then categorized into quintiles. For apartment building density and recent immigration status, the high frequency of zeros permitted the creation of only three categories (i.e., comparing the fourth and the fifth quintiles with the lower three quintiles).

*Statistical analysis*

The testing outcome was defined as receipt of at least one SARS-CoV-2 test during the study period. The comparator group comprised Ontario residents who had no record of testing during the study period. Determinants of testing were examined in unadjusted, age/sex-adjusted, and fully-adjusted logistic regression models that included all determinants. The fully-adjusted model also included a fixed-effect covariate for public health region. Public health regions are geographic areas in which public health measures were differentially applied[50] and along which there may be variability in measured and unmeasured social determinants.[51]

To address objective 1, we compared the odds of diagnosis via unadjusted, age/sex-adjusted, and fully-adjusted logistic regression models (including all determinants and public health regions) using three study designs. The “pseudo-test-negative design” compared individuals who tested positive to individuals who tested negative; the “true test-negative design” was restricted to tested individuals with symptomatic illness;[52] and

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

the "case-control" design compared all individuals who tested positive to all individuals who did *not* test positive (i.e., individuals who tested negative or were not tested).

To address objective 2, we focused on results from fully-adjusted logistic regression models from the pseudo-test-negative and case-control designs. We interpreted each set of determinants as independent analyses based on directed acyclic graphs (**Supplemental Figure 1**), with the case-control design having the least potential for collider bias based on our proposed causal diagrams.

Statistical analyses were conducted using SAS v.9.4 (Cary, NC). To assess for collinearity, we examined tolerances and variance inflation factors.

*Ethical review*

The use of data in this project was authorized under section 45 of Ontario's Personal Health Information Protection Act, which does not require review by a Research Ethics Board.

## Results

Of 758,691 individuals tested during the study period, 25,030 (3.3%) were diagnosed with COVID-19 (**Figure 1**). Only 11.8% of those tested had a symptom recorded by the provider, 13.6% were considered asymptomatic, and 74.6% were missing symptom information. Descriptive characteristics of the study population are reported in **Table 1** and **Supplemental Table 2**.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

*Determinants of testing*

In fully-adjusted analyses, odds of testing increased with age (**Table 2, Supplemental Table 3**). Males had lower odds of testing than females. Nearly every underlying health condition and most measures of prior healthcare use were associated with increased odds of testing. In contrast, higher ambient air pollution was associated with reduced odds of testing. There was little variability in the odds of testing by most area-based social determinants of health. However, areas with higher visible minority populations had lower odds of testing, whereas areas with higher household income and greater percentages of uncoupled individuals had higher odds of testing. Effect measures for most social determinants of health appeared to be progressively attenuated from unadjusted, to age/sex-adjusted, to fully-adjusted regression models (**Figure 2A, Supplemental Table 3**).

*Variability in determinants of diagnosis across analytic designs*

Comparison of analytic designs revealed an important influence on individual-level determinant effect measures, and less of an influence on social determinants (**Table 2**, **Supplemental Tables 4, 5, and 6)**. Variables associated with testing tended towards a variable relationship with diagnosis, depending on the analytic design. For example, the odds of testing for adults aged ≥85 years compared to those aged <5 years deviated considerably from the null (adjusted odds ratio [aOR]=5.60; 95%CI, 5.47-5.73), and the odds of diagnosis differed between the pseudo-test-negative design (aOR=1.76; 95%CI, 1.51-2.06) and the case-control design (aOR=7.26; 95%CI, 6.23-8.46) (**Table 2**). Some

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

health conditions associated with higher odds of testing, such as chronic respiratory conditions and indicators of prior healthcare use, appeared protective against diagnosis using the pseudo-test-negative design, but reverted to the null or showed increased odds of diagnosis using the case-control design. Results from the true test-negative design were largely similar to results from the pseudo-test-negative design with wider confidence intervals, with the exceptions that odds of diagnosis were higher for older ages using the true test-negative design compared to the pseudo-test-negative design, and lower for higher quintiles of essential workers in the true test-negative design compared to the pseudo-test-negative design (**Supplemental Tables 4 and 5**).

*Determinants of diagnosis using the case-control design*

Using the case-control design, older age, certain comorbidities (hypertension, diabetes, congestive heart failure, dementia, chronic kidney disease, and ischemic stroke/transient ischemic attack), and increased prior healthcare use were associated with increased odds of diagnosis. In contrast, other comorbidities (asthma, cancer, ischemic heart disease, substance abuse) and receipt of influenza vaccine in the 2019-2020 season were associated with reduced odds of diagnosis (**Table 2, Supplemental Table 6**).

The two highest categories of $PM_{2.5}$ exposure were associated with increased odds of diagnosis, whereas no categories of exposure to $NO_2$ were associated with increased odds of diagnosis.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Higher household density, increased apartment building density, greater percentages of uncoupled individuals, and greater percentages of essential workers were associated with greater risk of diagnosis. Lower educational attainment was related to an increased risk of diagnosis, but there was no statistically consistent relationship with household income. Finally, being in the highest quintile of neighbourhoods with visible minorities , and greater percentages of recent immigrants were associated with greater odds of diagnosis. Associations were attenuated after adjustment for all social determinants except for household density and essential work (**Figure 2B, Supplemental Table 6**).

In collinearity diagnostics, all tolerances were below 1 and all variance inflation factors were below 5 (**Supplemental Table 7**).

## Discussion

We identified variability in associations by analytic design, likely due to collider bias, across individual determinants of diagnosis. Using results from the case-control analysis, we identified particular individual, environmental, and social determinants of health as key determinants for diagnosis.

We found a high potential for erroneously identifying a “risk/protective factor” with individual determinants such as underlying health conditions, sex, and age. Our findings identified a pattern where underlying health conditions associated with COVID-19 severity[2] may have been prone to collider bias, where the direction of effect measures changed based on model choice for some conditions. This was further underlined by

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

higher testing and a seemingly protective effect when using the pseudo-test-negative design and true test-negative design for healthcare use variables that reversed when using the case-control design. Thus, comparisons of positivity rates require careful interpretation by examining the reasons for testing.[18] In the context of low overall levels of testing, the case-control design may have mitigated some potential sources of collider bias, with the assumption that those not tested are similar to those who tested negative.[17,18]

Some underlying health conditions remained associated with diagnoses using the case-control design, reflecting either unmeasured confounding or possible biological susceptibility to infection if exposed.[10,11,21,53,54] For example, dementia and frailty remained independently associated with diagnosis, potentially due to unmeasured confounding such as higher rates of contacts with care-givers or residence in other types of congregate settings such as retirement homes. Thus, underlying health conditions like dementia/frailty represent targets for prevention with strategies tailored to reduce exposures among individuals characterized by these individual determinants.

Over the study period, the testing criteria in Ontario shifted from returning symptomatic travellers, to severely symptomatic individuals and those with occupational exposure, to additional testing of asymptomatic individuals.[32-35,37] These distinctions may have created differences in symptomatic tested individuals, compared to all tested individuals. In our study, the restriction of the test-negative design to symptomatic individuals did not yield substantially different results than the test-negative design including

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

symptomatic and asymptomatic individuals for most determinants, but this may have been partly due to the high (74.6%) missingness of symptom information.

The independent association between high $PM_{2.5}$ and diagnosis may reflect unmeasured social determinants of health.[55,56] However, studies have also implicated environmental pollution as having a biological relationship to the risk and severity of COVID-19.[10,11,14]

We identified increased odds of diagnosis associated with household density, apartment building percentage, uncoupled status, essential work, educational attainment, and recent immigration, consistent with findings from other settings.[51,57,58] Household size has been shown to be a consistent risk factor across a broad range of settings.[59,60] These higher attack rates are likely due to prolonged and physically closer in-person contacts occurring more frequently within the household.[60] Essential services and occupations have also been associated with higher exposure risk,[61] either because such jobs cannot be done feasibly with proper protections, or because protective policies and materials are not issued, leaving workers at high risk.[62,63]

Higher percentages of recent immigrants in an area were associated with diagnosis, even after adjustment, although the percentage of visible minorities was not. Both variables might represent residual measures of structural racism, potentiating increased risk of SARS-CoV-2 exposure and COVID-19 severity,[64-66] including COVID-19-related hospitalization/death.[9,17,29,58] We found the association between visible minority status

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

and diagnosis was attenuated after adjustment for individual, environmental, and other social determinants of health. These findings likely reflect what is already known about race and ethnicity as social constructs and social determinants of health.[67] Finally, the fact that there was little variation in the odds of testing across levels of most of the social determinants suggests that testing resources may not be adequately prioritizing individuals at greatest risk.[68]

Taken together, our findings suggest a need to increase and/or redirect resources that specifically address social determinants such as household density[48,69] (e.g., voluntary isolation centres[70], wrap-around services[71]), occupational risk[62,66] (e.g., paid sick leave[72], workplace testing,[73] improved ventilation[62]), and other mediators of structural racism[68,74,75] (e.g. community-led outreach testing[76]). Our findings also suggest prioritizing COVID-19 vaccination strategies that reach communities and workplaces experiencing the highest rates of cases. Although there have been national recommendations for an equity lens to the public health response to COVID-19,[46] much of the existing response on COVID-19 equity and outreach to marginalized communities has been accomplished through smaller, independent groups, including volunteer organizations.[77-79]

**Limitations**

Our diagnoses were restricted to laboratory-confirmed cases and to the 88% of total provincial diagnoses that were available via OLIS, and thus may have missed probable cases as well as the remaining laboratory-confirmed cases that may have different

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

determinants of infection. Results are also conditioned on the assumption that determinants remained constant across the study period, whereas surveillance data suggest shifts in how infections propagate between social networks.[80] Future analyses include examining changes in the direction and magnitude of determinants over the course of the outbreak. Our models also adjusted for public health region, within which many social determinants cluster[51], and thus we cannot infer from the results presented how social determinants of diagnosis may vary between and within these geographic regions. Social determinants were measured at the area level and were not available at the individual level; however, by describing individuals' neighbourhoods, analyses reflect the role of structural and environmental determinants for individuals living amongst them. Fully-adjusted models in these analyses adjusted for a large number of covariates, potentially creating scenarios of over-adjustment. However, the directions of effect estimates generally remained the same after full adjustment, and the sample size of our analyses provided adequate statistical power. Finally, some important determinants, such as obesity,[23,81] were not available for our study.[82]

## Conclusion

Individual-level risks for diagnosis defined by demographic and health-related determinants representing general targets of current response strategies appear to be subject to collider bias. Moving forward and advancing the response necessitates characterizing and addressing the social determinants potentiating heterogeneity in acquisition and transmission risks with risk-tailored, community-based interventions to reduce COVID-19 burden.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Funding Statement**

This study was funded by the St. Michael's Hospital Research Innovation Council COVID-19 Research Grant, and research operating grant (VR5 -172683) from the Canadian Institutes of Health Research. SM is supported by a Tier 2 Canada Research Chair in Mathematical Modeling and Program Science. JCK is supported by a Clinician-Scientist Award from the University of Toronto Department of Family and Community Medicine. This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The study sponsors did not participate in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; or the decision to submit the manuscript for publication.

**Disclaimers:** The opinions, results, and conclusions reported in this paper are those of the authors and are independent from the funding sources. Parts of this material are based on data and/or information compiled and provided by the Canadian Institute for Health Information (CIHI) and by Cancer Care Ontario (CCO). However, the analyses, conclusions, opinions, and statement expressed herein are those of the authors, and not necessarily those of CIHI or CCO. No endorsement by ICES, MOHLTC, CIHI, or CCO is intended or should be inferred.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Acknowledgements:** We thank IQVIA Solutions Canada Inc. for use of their Drug Information File, as well as Owen Langman for technical assistance. Finally, we are grateful to the 14.7 million Ontario residents without whom this research would be impossible.

**Author contributions:**

JK, SM, and SB designed the study. ACa conducted all data analyses (dataset and variable creation and statistical modeling). JK, SM, SB, RK, ACa designed the analyses plans and conducted variable selection, with input from HC and AKC on variable selection and definitions. MAH, MD, LCR, TW contributed to analytic plans related to collider bias. BC contributed to data analyses and data preparation for the symptomatic dataset. MS prepared the figures. MS, JK, SB, and SM wrote the manuscript. All authors interpreted the data and critically reviewed and edited the manuscript.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

## References


1. Mishra S, Kwong JC, Chan AK, Baral SD. Understanding hetergeneity to inform public health response to COVID-19 in Canada. *Can Med Assoc J.* 2020;192(25):E684-E685.
2. *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19).* Geneva, Switzerland: World Health Organization; 16-24 February, 2020 2020.
3. Rosenberg ES, Dufort EM, Blog DS, et al. COVID-19 Testing, Epidemic Features, Hospital Outcomes, and Household Prevalence, New York State-March 2020. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2020.
4. Jing Q-L, Liu M-J, Yuan J, et al. Household secondary attack rate of COVID-19 and associated determinants. *medRxiv pre-print.* 2020.
5. Bhala N, Curry G, Martineau AR, Agyemang C, Bhopal R. Sharpening the global focus on ethnicity and race in the time of COVID-19. *The Lancet.*
6. Chung H, Fung K, Ferreira-Legere LE, et al. *COVID-19 Laboratory Testing in Ontario: Patterns of Testing and Characteristics of Individuals Tested, as of April 30, 2020.* ICES Ontario;2020.
7. Wang L, Mishra S. *The distribution of cumulative COVID-19 cases in Ontario by acquisition type over time.* 2020.
8. Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and Mortality among Black Patients and White Patients with Covid-19. *The New England journal of medicine.* 2020;382(26):2534-2543.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

9. Yehia BR, Winegar A, Fogel R, et al. Association of Race With Mortality Among Patients Hospitalized With Coronavirus Disease 2019 (COVID-19) at 92 US Hospitals. *JAMA Netw Open.* 2020;3(8):e2018039.

10. Feng C, Li J, Sun W, Zhang Y, Wang Q. Impact of ambient fine particulate matter (PM2.5) exposure on the risk of influenza-like-illness: a time-series analysis in Beijing, China. *Environ Health.* 2016;15:17.

11. Zoran MA, Savastru RS, Savastru DM, Tautan MN. Assessing the relationship between surface levels of PM2.5 and PM10 particulate matter impact on COVID-19 in Milan, Italy. *Sci Total Environ.* 2020;738:139825.

12. Wu X, Nethery RC, Sabath BM, Braun D, Dominici F. Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study. *medRxiv.* 2020.

13. Liang D, Shi L, Zhao J, et al. Urban Air Pollution May Enhance COVID-19 Case-Fatality and Mortality Rates in the United States. *medRxiv.* 2020.

14. Wu X, Nethery R, Sabath M, Braun D, Dominici F. Exposure to air pollution and COVID-19 mortality in the United States: a nationwide cross-sectional study. *medRxiv.* 2020.

15. Environmental racism: time to tackle social injustice. *The Lancet Planetary Health.* 2018;2(11).

16. Stall NM, Wu W, Lapointe-Shaw L, et al. Sex- and age-specific differences in COVID-19 testing, cases and outcomes: a population-wide study in Ontario, Canada. *J Am Geriatr Soc.* 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

17. Niedzwiedz CL, O'Donnell CA, Jani BD, et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. *BMC Med.* 2020;18(1):160.

18. Griffith G, Morris T, Tudball M, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *medRxiv.* 2020.

19. Hernán M, Robins J. *Causal Inference.* 2014.

20. Elwert F, Winship C. Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. *Annu Rev Sociol.* 2014;40:31-53.

21. Hussain A, Bhowmik B, do Vale Moreira NC. COVID-19 and diabetes: Knowledge in progress. *Diabetes Res Clin Pract.* 2020;162:108142.

22. Kumar A, Arora A, Sharma P, et al. Is diabetes mellitus associated with mortality and severity of COVID-19? A meta-analysis. *Diabetes Metab Syndr.* 2020;14(4):535-545.

23. Palaiodimos L, Kokkinidis DG, Li W, et al. Severe obesity, increasing age and male sex are independently associated with worse in-hospital outcomes, and higher in-hospital mortality, in a cohort of patients with COVID-19 in the Bronx, New York. *Metabolism.* 2020;108:154262.

24. Health care in Ontario: OHIP. https://www.ontario.ca/page/health-care-ontario. Published 2020. Accessed October 29, 2020.

25. *Schedule of benefits for laboratory services.* Ontario: Ministry of Health, Ontario Health Insurance Plan, Laboratories and Genetics Branch; July 1 2020.

26. IC/ES: Data, Discovery, Better Health. https://www.ices.on.ca. Published 2020. Accessed October 29, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

27. Data products, 2016 Census. Statistics Canada/Statistique Canada. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/index-eng.cfm. Published 2020. Accessed October 29, 2020.

28. Pennington E. Asthma increases risk of severity of COVID-19. *Cleve Clin J Med.* 2020.

29. Williamson EJ, Walker AJ, Bhaskaran K, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature.* 2020.

30. Fang L, Karakiulakis G, Roth M. Are patients with hyperetension and diabetes mellitus at increased risk for COVID-19 infection? *Lancet Respiratory Medicine.* 2020;8:e21.

31. Merkler AE, Parikh NS, Mir S, et al. Risk of ischemic stroke in patients with COVID-19 versus patients with influenza. *medRxiv.* 2020.

32. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: May 14, 2020.* May 14 2020.

33. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: May 2, 2020.* May 2 2020.

34. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: April 15, 2020.* April 15 2020.

35. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: May 28, 2020.* May 28 2020.

36. Ontario Ministry of Health. *COVID-19 Quick Reference Public Health Guidance on Testing and Clearance: March 27, 2020.* 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

37. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: April 8, 2020.* April 8, 2020 2020.

38. Tolea MI, Morris JC, Galvin JE. Trajectory of Mobility Decline by Type of Dementia. *Alzheimer Dis Assoc Disord.* 2016;30(1):60-66.

39. Chiao C-Y, Wu H-S, Hsiao C-Y. Caregiver burden for informal caregivers of patients with dementia: A systematic review. *Int Nurs Rev.* 2015;62(3):340-350.

40. Johns Hopkins ACG(c) System. Johns Hopkins. https://www.hopkinsacg.org. Published 2020. Accessed November 1, 2020.

41. Austin PC, van Walraven C, Wodchis WP, Newman A, Anderson GM. Using the Johns Hopkins Aggregated Diagnosis Groups (ADGs) to predict mortality in a general adult population cohort in Ontario, Canada. *Med Care.* 2011;49(10):932-939.

42. van Donkelaar A, Martin RV, Brauer M, Boys BL. Use of Satellite Observations for Long-Term Exposure Assessment of Global Concentrations of Fine Particulate Matter. *Environ Health Perspect.* 2015;123:135-143.

43. Hystad P, Setton E, Cervantes A, et al. Creating national air pollution models for population exposure assessment in Canada. *Environ Health Perspect.* 2011;119(8):1123-1129.

44. Walker P, Whittaker C, Watson O, et al. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science.* 2020;369(6502):413-422.

45. Park SY, Kim YM, Yi S, et al. Coronavirus Disease Outbreak in Call Center, South Korea. *Emerg Infect Dis.* 2020;26(8):1666-1670.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

46. *From risk to resilience: an equity approach to COVID-19.* Ottawa, Canada: Office of the Chief Public Health Officer of Canada;2020.

47. Karaye IM, Horney JA. The Impact of Social Vulnerability on COVID-19 in the U.S.: An Analysis of Spatially Varying Relationships. *American journal of preventive medicine.* 2020;59(3):317-325.

48. Rodriguez-Diaz CE, Guilamo-Ramos V, Mena L, et al. Risk for COVID-19 infection and death among Latinos in the United States: examining heterogeneity in transmission dynamics. *Ann Epidemiol.* 2020.

49. Statistics Canada/Statistique Canada. Dictionary, Census of Population, 2016. https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/geo021-eng.cfm. Published 2020. Accessed October 13, 2020.

50. Public Health Units. https://www.ontario.ca/page/public-health-units. Published 2017. Updated March 22, 2019. Accessed October 29, 2020.

51. *Enhanced epidemiological summary: COVID-19 in Ontario—a focus on diversity.* Public Health Ontario;2020.

52. Sullivan SG, Tchetgen Tchetgen EJ, Cowling BJ. Theoretical Basis of the Test-Negative Study Design for Assessment of Influenza Vaccine Effectiveness. *American journal of epidemiology.* 2016;184(5):345-353.

53. Nishiga M, Wang DW, Han Y, Lewis DB, Wu JC. COVID-19 and cardiovascular disease: from basic mechanisms to clinical perspectives. *Nat Rev Cardiol.* 2020;17(9):543-558.

54. Elliot AJ, Harcourt SE, Hughes HE, et al. The COVID-19 pandemic: a new challenge for syndromic surveillance. *Epidemiol Infect.* 2020;148:e122.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

55. Morelli X, Rieux C, Cyrys J, Forsberg B, Slama R. Air pollution, health and social deprivation: A fine-scale risk assessment. *Environ Res.* 2016;147:59-70.

56. Næss Ø, Piro FN, Nafstad P, Smith GD, Leyland AH. Air Pollution, Social Deprivation, and Mortality. *Epidemiology.* 2007;18(6):686-694.

57. Lin SL. Intersectionality and inequalities in medical risk for severe COVID-19 in the Canadian Longitudinal Study on Aging. *Gerontologist.* 2020.

58. Drefahl S, Wallace M, Mussino E, et al. A population-based cohort study of socio-demographic risk factors for COVID-19 deaths in Sweden. *Nat Commun.* 2020;11(1):5097.

59. Luo L, Liu D, Liao X, et al. Contact Settings and Risk for Transmission in 3410 Close Contacts of Patients With COVID-19 in Guangzhou, China : A Prospective Cohort Study. *Ann Intern Med.* 2020;173(11):879-887.

60. Jing Q-L, Liu M-J, Zhang Z-B, et al. Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study. *The Lancet Infectious Diseases.* 2020;20(10):1141-1150.

61. *Workplace Safety & Prevention Services Guidance on Health and Safety for the Personal Care Service Sector during COVID-19.* Ontario: Workplace Safety & Prevention Services;2020.

62. Dyal J, Grant M, Broadwater K, et al. COVID-19 among workers in meat and poultry processing facilities—19 states, April 2020. *Morbidity and Mortality Weekly Report.* 2020;69(18):557-561.

63. Contrera J. The N95 shortage America can't seem to fix. *Washington Post.* September 21, 2020, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

64. Vahidy FS, Nicolas JC, Meeks JR, et al. Racial and ethnic disparities in SARS-CoV-2 pandemic: analysis of a COVID-19 observational registry for a diverse US metropolitan population. *BMJ Open.* 2020;10(8):e039849.

65. de Lusignan S, Dorward J, Correa A, et al. Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of General Practitioners Research and Surveillance Centre primary care network: a cross-sectional study. *The Lancet Infectious Diseases.* 2020;20(9):1034-1042.

66. Liem A, Wang C, Wariyanti Y, Latkin CA, Hall BJ. The neglected health of international migrant workers in the COVID-19 epidemic. *The Lancet Psychiatry.* 2020;7(4).

67. Ordóñez CE. Not race, racism: Concerns of COVID-19 affecting African Americans. *Anthropological Responses to Heatlh Emergencies.* 2020.

68. Dodds C, Fakoya I. Covid-19: ensuring equality of access to testing for ethnic minorities. *BMJ.* 2020;369:m2122.

69. Brandén M, Aradhya S, Kolk M, et al. Residential context and COVID-19 mortality among adults aged 70 years and older in Stockholm: a population-based, observational study using individual-level data. *The Lancet Healthy Longevity.* 2020.

70. Jordan-Martin NC, Madad S, Alves L, et al. Isolation Hotels: A Community-Based Intervention to Mitigate the Spread of the COVID-19 Pandemic. *Health Secur.* 2020.

71. Madad S, Nuzzo J, Bordeaux M. The missing piece in America's COVID-19 isolation and quarantine strategy: Wraparound services.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

https://www.healthaffairs.org/do/10.1377/hblog20201207.458415/full/. Published 2020. Accessed March 9, 2021.

72. Vazquez J, Islam T, Beller J, et al. Expanding paid sick leave as a public health tool in the COVID-19 pandemic. *J Occup Environ Med.* 2020;Letter to the Editor.

73. Rapid COVID-19 testing a potential game-changer in worker protection. Institute for Work and Health. https://www.iwh.on.ca/newsletters/at-work/102/rapid-covid-19-testing-potential-game-changer-in-worker-protection. Published 2020. Accessed March 9, 2021.

74. Webb Hooper M, Napoles AM, Perez-Stable EJ. COVID-19 and Racial/Ethnic Disparities. *JAMA.* 2020.

75. Yang TC, Emily Choi SW, Sun F. COVID-19 cases in US counties: roles of racial/ethnic density and residential segregation. *Ethn Health.* 2020:1-11.

76. Wallis G, Siracusa F, Blank M, et al. Experience of a novel community testing programme for COVID-19 in London: Lessons learnt. *Clin Med (Lond).* 2020;20(5):e165-e169.

77. Home | Good Neighbour Project. https://www.goodneighbourproject.com. Published 2021. Accessed February 22, 2021.

78. Marwaha S, Vohra-Miller S, Grewal R. We started the South Asian COVID Task Force because Ontario failed to address inequities. In a short time, we've seen more people get tested. *Toronto Star.* December 17, 2020, 2020.

79. Jabakhanji S. Meet the Toronto Indigenous organizations bringing COVID-19 testing, food directly to people's doorsteps. *CBC.* December 19, 2020, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

80. Wang L, Ma H, Yiu KCY, et al. Heterogeneity in risk, testing and outcome of COVID-19 across outbreak settings in the Greater Toronto Area, Canada: an observational study. *medRxiv.* 2020.

81. Kwok S, Adam S, Ho JH, et al. Obesity: A critical risk factor in the COVID-19 pandemic. *Clin Obes.* 2020:e12403.

82. ICES Data Dictionary. ICES. https://datadictionary.ices.on.ca/Applications/DataDictionary/Default.aspx. Published 2020. Accessed November 1, 2020.

**Table 1**. Characteristics of study population tested and not tested for SARS-CoV-2 in Ontario (March 1, 2020 to June 20, 2020).

| | **Tested individuals** | | **All individuals** | |
|---|---|---|---|---|
| | **Tested negative (N = 733,661)** | **Tested positive (N = 25,030)** | **Untested control group (N = 13,936,888)** | **All tested (N =758,691)** |
| **Demographic characteristics** | | | | |
| Age: mean ± SD | 49.52 ± 20.99 | 47.45 ± 19.77 | 40.31 ± 22.91 | 49.45 ± 20.95 |
| Male sex: N (%) | 289,769 (39.5%) | 11,627 (46.5%) | 6,930,233 (49.7%) | 301,396 (39.7%) |
| Living in rural/small town: N (%)[a] | 82,034 (11.2%) | 887 (3.5%) | 1,390,426 (10.0%) | 82,921 (10.9%) |
| **Chronic health conditions: N (%)** | | | | |
| Asthma | 136,556 (18.6%) | 3,887 (15.5%) | 2,085,118 (15.0%) | 140,443 (18.5%) |
| COPD | 34,261 (4.7%) | 715 (2.9%) | 249,636 (1.8%) | 34,976 (4.6%) |
| Hypertension | 226,111 (30.8%) | 7,488 (29.9%) | 2,836,790 (20.4%) | 233,599 (30.8%) |
| Diabetes | 108,758 (14.8%) | 4,268 (17.1%) | 1,363,667 (9.8%) | 113,026 (14.9%) |
| Congestive heart failure | 35,947 (4.9%) | 983 (3.9%) | 225,153 (1.6%) | 36,930 (4.9%) |
| Dementia or frailty | 31,804 (4.3%) | 1,106 (4.4%) | 124,295 (0.9%) | 32,910 (4.3%) |
| Cancer[b] | 24,333 (3.3%) | 484 (1.9%) | 213,036 (1.5%) | 24,817 (3.3%) |
| Chronic kidney disease[b] | 37,630 (5.1%) | 1,122 (4.5%) | 254,541 (1.8%) | 38,752 (5.1%) |
| Immunocompromised[c] | 13,325 (1.8%) | 319 (1.3%) | 107,824 (0.8%) | 13,644 (1.8%) |
| Advanced liver disease | 9,020 (1.2%) | 212 (0.8%) | 78,747 (0.6%) | 9,232 (1.2%) |
| Ischemic heart disease | 33,855 (4.6%) | 880 (3.5%) | 324,682 (2.3%) | 34,735 (4.6%) |
| Ischemic stroke or transient ischemic attack[d] | 13,777 (1.9%) | 411 (1.6%) | 99,503 (0.7%) | 14,188 (1.9%) |
| Schizophrenia[e] | 7,526 (1.0%) | 220 (0.9%) | 64,494 (0.5%) | 7,746 (1.0%) |
| Substance abuse[e] | 22,238 (3.0%) | 478 (1.9%) | 216,156 (1.6%) | 22,716 (3.0%) |
| **Healthcare use: N (%)** | | | | |
| **Adjusted Diagnostic Group quintile** | | | | |
| 1 (0 ADGs) | 31,959 (4.4%) | 1,571 (6.3%) | 1,616,852 (11.6%) | 33,530 (4.4%) |
| 2 (1-2 ADGs) | 101,895 (13.9%) | 3,601 (14.4%) | 3,150,969 (22.6%) | 105,496 (13.9%) |
| 3 (3-4 ADGs) | 149,173 (20.3%) | 5,101 (20.4%) | 3,385,602 (24.3%) | 154,274 (20.3%) |
| 4 (5-6 ADGs) | 150,394 (20.5%) | 5,115 (20.4%) | 2,626,403 (18.8%) | 155,509 (20.5%) |
| 5 (7-27 ADGs) | 300,240 (40.9%) | 9,642 (38.5%) | 3,157,062 (22.7%) | 309,882 (40.8%) |
| Hospital admissions, past 3 years: Mean (SD) | 0.44 ± 1.34 | 0.33 ± 1.21 | 0.19 ± 0.60 | 0.44 ± 1.34 |
| Outpatient physician visits, past | 7.35 ± 8.68 | 7.06 ± 8.33 | 4.67 ± 6.05 | 7.34 ± 8.67 |



medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

| | | | | |
|---|---|---|---|---|
| year: Mean (SD) | | | | |
| Influenza vaccination (2019-2020 season) | 213,722 (29.1%) | 5,547 (22.2%) | 2,978,472 (21.4%) | 219,269 (28.9%) |
| **Environmental determinants[f]: N (%)** | | | | |
| **$PM_{2.5}$ (μg/m$^3$ yearly)** | | | | |
| 2 to <6 | 161,300 (22.0%) | 1,831 (7.3%) | 2,481,201 (17.8%) | 163,131 (21.5%) |
| 6 to <7 | 91,134 (12.4%) | 1,766 (7.1%) | 1,555,790 (11.2%) | 92,900 (12.2%) |
| 7 to <8 | 207,966 (28.3%) | 8,476 (33.9%) | 4,450,218 (31.9%) | 216,442 (28.5%) |
| 8 to <9 | 211,861 (28.9%) | 11,127 (44.5%) | 4,258,069 (30.6%) | 222,988 (29.4%) |
| ≥9 | 59,674 (8.1%) | 1,747 (7.0%) | 1,073,665 (7.7%) | 61,421 (8.1%) |
| **$NO_2$ (ppb yearly)** | | | | |
| 0 to 6 | 328,613 (44.8%) | 5,237 (20.9%) | 5,505,976 (39.5%) | 333,850 (44.0%) |
| 6 to 8 | 170,693 (23.3%) | 5,599 (22.4%) | 3,409,506 (24.5%) | 176,292 (23.2%) |
| ≥8 | 232,629 (31.7%) | 14,111 (56.4%) | 4,903,460 (35.2%) | 246,740 (32.5%) |
| **Social determinants of health[g] (area-level): N (%)** | | | | |
| **Household density quintile[h]** | | | | |
| 1 (0 – 2.1) | 162,623 (22.2%) | 3,639 (14.5%) | 2,474,391 (17.8%) | 166,262 (21.9%) |
| 2 (2.2 - 2.4) | 140,653 (19.2%) | 3,104 (12.4%) | 2,368,013 (17.0%) | 143,757 (18.9%) |
| 3 (2.5 - 2.6) | 104,896 (14.3%) | 2,721 (10.9%) | 1,866,317 (13.4%) | 107,617 (14.2%) |
| 4 (2.7 - 3) | 166,089 (22.6%) | 6,321 (25.3%) | 3,291,097 (23.6%) | 172,410 (22.7%) |
| 5 (3.1 - 5.7) | 152,578 (20.8%) | 8,929 (35.7%) | 3,775,489 (27.1%) | 161,507 (21.3%) |
| **Apartment building density category[i]** | | | | |
| 1 (0 – 7.3%) | 391,477 (53.4%) | 12,377 (49.4%) | 7,994,323 (57.4%) | 403,854 (53.2%) |
| 2 (7.4 - 37.7%) | 145,108 (19.8%) | 3,874 (15.5%) | 2,478,855 (17.8%) | 148,982 (19.6%) |
| 3 (37.7 - 104%) | 190,244 (25.9%) | 8,463 (33.8%) | 3,301,839 (23.7%) | 198,707 (26.2%) |
| **Uncoupled quintile[j]** | | | | |
| 1 (11.2 - 33.7%) | 150,044 (20.5%) | 3,539 (14.1%) | 3,130,907 (22.5%) | 153,583 (20.2%) |
| 2 (33.7 - 38.4%) | 128,561 (17.5%) | 3,915 (15.6%) | 2,698,847 (19.4%) | 132,476 (17.5%) |
| 3 (38.5 - 43.6%) | 127,689 (17.4%) | 4,632 (18.5%) | 2,579,005 (18.5%) | 132,321 (17.4%) |
| 4 (43.6 - 51.0%) | 145,560 (19.8%) | 5,591 (22.3%) | 2,633,710 (18.9%) | 151,151 (19.9%) |
| 5 (51.0 - 94.6%) | 174,985 (23.9%) | 7,037 (28.1%) | 2,732,838 (19.6%) | 182,022 (24.0%) |
| **Essential worker status[k]** | | | | |
| 1 (0 – 32.5%) | 145,517 (19.8%) | 3,941 (15.7%) | 2,969,492 (21.3%) | 149,458 (19.7%) |
| 2 (32.5 – 42.3%) | 155,189 (21.2%) | 5,077 (20.3%) | 3,084,781 (22.1%) | 160,266 (21.1%) |
| 3 (42.3 – 49.8%) | 149,589 (20.4%) | 4,686 (18.7%) | 2,717,510 (19.5%) | 154,275 (20.3%) |
| 4 (50.0 – 57.5%) | 143,965 (19.6%) | 5,236 (20.9%) | 2,615,078 (18.8%) | 149,201 (19.7%) |



medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

| | | | | |
|---|---|---|---|---|
| 5 (57.5 – 114.3%) | 133,601 (18.2%) | 5,816 (23.2%) | 2,399,120 (17.2%) | 139,417 (18.4%) |
| **Household income quintile[l]** | | | | |
| 1 (lowest income) | 156,320 (21.3%) | 7,000 (28.0%) | 2,679,780 (19.2%) | 163,320 (21.5%) |
| 2 | 148,687 (20.3%) | 5,288 (21.1%) | 2,698,807 (19.4%) | 153,975 (20.3%) |
| 3 | 145,317 (19.8%) | 5,084 (20.3%) | 2,791,340 (20.0%) | 150,401 (19.8%) |
| 4 | 140,352 (19.1%) | 4,019 (16.1%) | 2,809,529 (20.2%) | 144,371 (19.0%) |
| 5 (highest income) | 138,103 (18.8%) | 3,419 (13.7%) | 2,815,238 (20.2%) | 141,522 (18.7%) |
| **Limited educational attainment quintile[m]** | | | | |
| 1 (0.0 - 4.1%) | 144,457 (19.7%) | 3,903 (15.6%) | 2,926,432 (21.0%) | 148,360 (19.6%) |
| 2 (4.1 - 7.5%) | 154,215 (21.0%) | 4,477 (17.9%) | 2,978,444 (21.4%) | 158,692 (20.9%) |
| 3 (7.5 - 11.4%) | 151,457 (20.6%) | 5,052 (20.2%) | 2,888,525 (20.7%) | 156,509 (20.6%) |
| 4 (11.4 - 17.1%) | 148,159 (20.2%) | 5,314 (21.2%) | 2,659,544 (19.1%) | 153,473 (20.2%) |
| 5 (17.1 - 94.3%) | 129,580 (17.7%) | 6,010 (24.0%) | 2,333,134 (16.7%) | 135,590 (17.9%) |
| **Visible minority quintile[n]** | | | | |
| 1 (0.0 - 2.2%) | 130,912 (17.8%) | 1,716 (6.9%) | 2,115,641 (15.2%) | 132,628 (17.5%) |
| 2 (2.2 - 7.5%) | 137,826 (18.8%) | 2,233 (8.9%) | 2,255,245 (16.2%) | 140,059 (18.5%) |
| 3 (7.5 - 18.7%) | 137,744 (18.8%) | 3,201 (12.8%) | 2,451,335 (17.6%) | 140,945 (18.6%) |
| 4 (18.7 - 43.5%) | 153,503 (20.9%) | 5,466 (21.8%) | 3,023,752 (21.7%) | 158,969 (21.0%) |
| 5 (43.5 - 102%) | 167,893 (22.9%) | 12,140 (48.5%) | 3,940,245 (28.3%) | 180,033 (23.7%) |
| **Recent immigration category[o]** | | | | |
| 1 (0.0 - 2.1%) | 401,300 (54.7%) | 8,271 (33.0%) | 6,967,468 (50.0%) | 409,571 (54.0%) |
| 2 (2.1 - 4.7% | 146,772 (20.0%) | 5,409 (21.6%) | 2,858,064 (20.5%) | 152,181 (20.1%) |
| 3 (4.7 - 41.2%) | 174,829 (23.8%) | 11,018 (44.0%) | 3,898,463 (28.0%) | 185,847 (24.5%) |

[a] "Rural" defined as located outside the commuting zone of a city with greater than 10,000 population.
[b] Participants were counted if they had a diagnosis in the last 5 years.
[c] Participants considered "immunocompromised" if they were HIV positive, or had an organ or bone marrow transplant, or another immunodeficient condition.
[d] This category includes individuals with ischemic stroke or TIA in the last 20 years.
[e] This category includes individuals with a diagnosis in the last 2 years.
[f] Values of $PM_{2.5}$ > 12µg/m$^3$ or $NO_2$ > 53ppb yearly has been found to be associated with increased risk of other respiratory illnesses.[83] $PM_{2.5}$ and $NO_2$ values are provided at the postal code, not DA, level.
[g] All variables in this category are area-level variables at the level of the Census Dissemination Area (DA).
[h] Range of persons per dwelling.
[i] The census counts are randomly rounded up or down to the nearest number divisible by 5, which causes some minor imprecision. 7.3% represents the 60$^{th}$ percentile.



medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[j] "Uncoupled" individuals are those never married (persons who have never legally married and are not living with a person as a couple); separated (persons who are married but who are no longer living with their spouse [for reasons other than, for example, illness, work or school], have not obtained a divorce and are not living with a person as a couple; divorced (persons who have obtained a legal divorce, have not remarried and are not living with a person as a couple); and widowed (persons who have lost their married spouse through death, have not remarried and are not living with a person as a couple).

[k] Percentage of individuals in the area working in the following occupations: Sales and service occupations; trades, transport and equipment operators and related occupations; natural resources, agriculture and related production occupations; and occupations in manufacturing and utilities. Census counts for individuals are randomly rounded up or down to the nearest number divisible by 5, which causes some minor imprecision.

[l] Income quintile has variable cut-off values in each city or census area, to take cost of living into account. A DA being in quintile 1 means it is among the lowest 20% DAs in its city by income.

[m] Percentage of adults aged 25-64 in the area not having received any diploma.

[n] Percentage of individuals in the area self-identifying as a visible minority. Census counts for individuals are randomly rounded up or down to the nearest number divisible by 5, which causes some minor imprecision.

[o] Percentage of individuals in the area who are recent immigrants; 2.1% represents the 60$^{th}$ percentile.



medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Table 2.** Odds of ever being tested for SARS-CoV-2 and of COVID-19 diagnosis in fully-adjusted analyses, using two analytic designs, in Ontario, Canada between March 1 and June 20, 2020.

| | **Tested for SARS-CoV-2** | **Diagnosed with COVID-19 [a]** | |
|---|---|---|---|
| **Determinants** | **Tested vs. not tested aOR (95% CI)** | **Test-positive vs. test-negative among all individuals tested (Pseudo-test-negative design) aOR (95% CI)** | **Test-positive vs. not test-positive (i.e., all negatives and all untested individuals) (Case-control design) aOR (95% CI)** |
| **Sample size** | 14,695,579 | 758,691 | 14,695,579 |
| **Individual determinants** | | | |
| **Age groups (years) (ref: 0 – 4)** | | | |
| 5-19 | 1.15 (1.13, 1.18) | 1.93 (1.66, 2.24) | 1.72 (1.49, 1.99) |
| 20-34 | 4.13 (4.05, 4.22) | 2.00 (1.75, 2.30) | 7.03 (6.13, 8.05) |
| 35-49 | 4.13 (4.05, 4.21) | 1.98 (1.73, 2.27) | 6.74 (5.88, 7.73) |
| 50-64 | 4.21 (4.12, 4.29) | 2.02 (1.76, 2.32) | 7.01 (6.11, 8.03) |
| 65-74 | 3.04 (2.98, 3.11) | 1.62 (1.40, 1.87) | 3.99 (3.46, 4.60) |
| 75-84 | 3.01 (2.95, 3.08) | 1.64 (1.41, 1.90) | 3.94 (3.39, 4.57) |
| ≥85 | 5.60 (5.47, 5.73) | 1.76 (1.51, 2.06) | 7.26 (6.23, 8.46) |
| Male sex | 0.76 (0.76, 0.76) | 1.26 (1.23, 1.30) | 1.00 (0.98, 1.03) |
| Living in rural/small town[b] | 0.97 (0.96, 0.98) | 0.81 (0.74, 0.88) | 0.76 (0.70, 0.82) |
| **Underlying chronic health conditions** | | | |
| Asthma | 1.09 (1.09, 1.10) | 0.86 (0.83, 0.89) | 0.92 (0.89, 0.96) |
| COPD | 1.23 (1.21, 1.24) | 0.89 (0.82, 0.96) | 1.04 (0.96, 1.13) |
| Hypertension | 0.98 (0.98, 0.99) | 1.12 (1.08, 1.16) | 1.13 (1.09, 1.17) |
| Diabetes | 0.98 (0.97, 0.99) | 1.26 (1.21, 1.31) | 1.19 (1.14, 1.23) |
| Congestive heart failure | 1.26 (1.24, 1.28) | 0.99 (0.92, 1.07) | 1.25 (1.16, 1.35) |
| Dementia or frailty score >15 | 2.12 (2.09, 2.15) | 1.39 (1.29, 1.49) | 2.60 (2.42, 2.80) |
| Cancer[c] | 1.13 (1.12, 1.15) | 0.72 (0.66, 0.80) | 0.86 (0.78, 0.94) |
| Chronic kidney disease[c] | 1.31 (1.29, 1.32) | 0.86 (0.80, 0.93) | 1.16 (1.08, 1.24) |
| Immunocompromised[d] | 1.30 (1.28, 1.33) | 0.79 (0.70, 0.89) | 1.02 (0.91, 1.14) |
| Advanced liver disease | 1.14 (1.11, 1.16) | 0.79 (0.68, 0.90) | 0.87 (0.76, 1.00) |
| Ischemic heart disease | 1.02 (1.01, 1.03) | 0.88 (0.82, 0.95) | 0.90 (0.84, 0.97) |
| Ischemic stroke or transient ischemic attack[e] | 1.15 (1.13, 1.17) | 1.05 (0.95, 1.17) | 1.18 (1.06, 1.31) |
| Schizophrenia[f] | 1.24 (1.21, 1.27) | 0.84 (0.73, 0.97) | 0.99 (0.87, 1.13) |



medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

| | | | |
|---|---|---|---|
| Substance abuse[f] | 1.17 (1.16, 1.19) | 0.72 (0.66, 0.79) | 0.86 (0.79, 0.95) |
| **Healthcare use** | | | |
| **Aggregated Diagnostic Group quintile (ref: 0 ADGs)** | | | |
| 2 (1-2 ADGs) | 1.64 (1.62, 1.66) | 0.78 (0.73, 0.84) | 1.24 (1.17, 1.32) |
| 3 (3-4 ADGs) | 2.07 (2.04, 2.10) | 0.75 (0.70, 0.80) | 1.46 (1.37, 1.56) |
| 4 (5-6 ADGs) | 2.44 (2.41, 2.48) | 0.74 (0.69, 0.79) | 1.69 (1.57, 1.81) |
| 5 (7-27 ADGs) | 3.11 (3.06, 3.15) | 0.75 (0.69, 0.80) | 2.12 (1.97, 2.28) |
| **Hospital admissions, past 3 years (ref: 0 admissions)** | | | |
| 1 | 0.99 (0.98, 1.00) | 0.93 (0.89, 0.97) | 0.92 (0.88, 0.96) |
| 2 | 1.16 (1.14, 1.17) | 0.86 (0.79, 0.93) | 1.01 (0.94, 1.09) |
| ≥3 | 1.75 (1.73, 1.78) | 0.82 (0.76, 0.90) | 1.38 (1.27, 1.50) |
| **Outpatient physician visits, past year (ref: 0 – 1 visits)** | | | |
| 2 – 4 | 1.08 (1.07, 1.09) | 1.02 (0.98, 1.07) | 1.12 (1.08, 1.18) |
| 5 – 8 | 1.11 (1.10, 1.12) | 1.03 (0.98, 1.08) | 1.17 (1.11, 1.23) |
| 9 – 14 | 1.17 (1.16, 1.18) | 0.98 (0.93, 1.04) | 1.17 (1.10, 1.24) |
| ≥15 | 1.36 (1.34, 1.37) | 0.91 (0.85, 0.97) | 1.22 (1.14, 1.30) |
| Influenza vaccination (2019-2020 season) | 1.08 (1.08, 1.09) | 0.81 (0.78, 0.83) | 0.87 (0.85, 0.90) |
| **Environmental determinants[g]** | | | |
| **$PM_{2.5}$ category (μg/m$^3$) (ref: 2 – 6)** | | | |
| 6-7 | 0.97 (0.97, 0.98) | 0.91 (0.85, 0.99) | 0.92 (0.85, 0.99) |
| 7-8 | 0.92 (0.90, 0.93) | 1.10 (0.99, 1.21) | 1.00 (0.91, 1.10) |
| 8-9 | 0.91 (0.90, 0.93) | 1.29 (1.16, 1.43) | 1.19 (1.08, 1.32) |
| ≥10 | 0.90 (0.89, 0.92) | 1.45 (1.29, 1.63) | 1.31 (1.16, 1.47) |
| **$NO_2$ category (parts per billion) (ref: 0 – 6)** | | | |
| 6-8 | 0.95 (0.95, 0.96) | 1.05 (1.00, 1.11) | 1.00 (0.96, 1.06) |
| ≥8 | 0.94 (0.93, 0.95) | 1.13 (1.06, 1.21) | 1.05 (0.98, 1.12) |
| **Social determinants of health[h] (area-level)** | | | |
| **Household density quintile[i] (ref: 1st quintile)** | | | |
| 2 | 1.01 (1.00, 1.02) | 1.21 (1.14, 1.27) | 1.19 (1.13, 1.26) |
| 3 | 1.03 (1.02, 1.04) | 1.39 (1.31, 1.48) | 1.42 (1.34, 1.50) |
| 4 | 1.01 (1.00, 1.02) | 1.70 (1.61, 1.79) | 1.70 (1.61, 1.80) |
| 5 | 0.97 (0.95, 0.98) | 1.94 (1.82, 2.07) | 1.86 (1.75, 1.98) |
| **Apartment building density category[j] (ref: 1st category)** | | | |
| 2 | 1.04 (1.03, 1.04) | 1.00 (0.95, 1.04) | 1.02 (0.98, 1.06) |
| 3 | 1.01 (1.00, 1.02) | 1.15 (1.09, 1.21) | 1.18 (1.12, 1.24) |
| **Uncoupled quintile[k] (ref: 1st quintile)** | | | |



medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

| | | | |
|---|---|---|---|
| 2 | 1.02 (1.01, 1.03) | 0.96 (0.91, 1.01) | 0.97 (0.92, 1.02) |
| 3 | 1.07 (1.06, 1.08) | 0.95 (0.90, 1.00) | 0.99 (0.94, 1.04) |
| 4 | 1.19 (1.18, 1.21) | 0.96 (0.91, 1.02) | 1.11 (1.05, 1.17) |
| 5 | 1.39 (1.38, 1.41) | 1.07 (1.01, 1.15) | 1.41 (1.32, 1.51) |
| **Essential work quintile[l] (ref: 1st category)** | | | |
| 2 | 1.04 (1.03, 1.04) | 1.25 (1.19, 1.32) | 1.30 (1.24, 1.37) |
| 3 | 1.06 (1.05, 1.07) | 1.28 (1.21, 1.35) | 1.37 (1.30, 1.45) |
| 4 | 1.05 (1.04, 1.06) | 1.37 (1.29, 1.45) | 1.51 (1.42, 1.60) |
| 5 | 1.04 (1.03, 1.06) | 1.42 (1.32, 1.51) | 1.58 (1.48, 1.69) |
| **Household income quintile (ref: 1st quintile, lowest income)[m]** | | | |
| 2 | 1.03 (1.02, 1.04) | 0.96 (0.91, 1.00) | 1.00 (0.96, 1.05) |
| 3 | 1.08 (1.07, 1.09) | 1.02 (0.97, 1.08) | 1.12 (1.06, 1.18) |
| 4 | 1.10 (1.09, 1.11) | 0.97 (0.90, 1.03) | 1.06 (1.00, 1.13) |
| 5 | 1.11 (1.09, 1.12) | 0.97 (0.90, 1.04) | 1.07 (0.99, 1.15) |
| **Limited educational attainment quintile[n] (ref: 1st quintile)** | | | |
| 2 | 1.03 (1.02, 1.04) | 1.07 (1.02, 1.12) | 1.09 (1.04, 1.15) |
| 3 | 1.01 (1.00, 1.02) | 1.15 (1.09, 1.21) | 1.15 (1.09, 1.21) |
| 4 | 1.02 (1.01, 1.03) | 1.23 (1.17, 1.30) | 1.21 (1.15, 1.28) |
| 5 | 1.01 (1.00, 1.02) | 1.37 (1.29, 1.46) | 1.33 (1.26, 1.41) |
| **Visible minority quintile[o] (ref: 1st category)** | | | |
| 2 | 0.99 (0.98, 1.00) | 0.96 (0.89, 1.02) | 0.95 (0.89, 1.02) |
| 3 | 0.95 (0.94, 0.96) | 0.97 (0.91, 1.05) | 0.93 (0.87, 1.00) |
| 4 | 0.91 (0.90, 0.92) | 1.07 (0.99, 1.15) | 0.98 (0.91, 1.06) |
| 5 | 0.86 (0.85, 0.87) | 1.27 (1.17, 1.38) | 1.09 (1.00, 1.19) |
| **Recent immigration category[p] (ref: 1st category)** | | | |
| 2 | 0.99 (0.99, 1.00) | 1.04 (1.00, 1.09) | 1.04 (1.00, 1.08) |
| 3 | 0.94 (0.93, 0.95) | 1.16 (1.11, 1.22) | 1.10 (1.05, 1.15) |

[a] All models presented are fully-adjusted models, containing all variables listed in this table as covariates.
[b] "Rural" defined as located outside the commuting zone of a city with greater than 10,000 population.
[c] Participants were counted if they had a diagnosis in the last 5 years.
[d] Participants considered "immunocompromised" if they were HIV positive, or had an organ or bone marrow transplant, or another immunodeficient condition.
[e] This category includes individuals with ischemic stroke or transient ischemic attack in the last 20 years.
[f] This category includes individuals with a diagnosis in the last 2 years.



medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[g] Values of $PM_{2.5}$ > 12µg/$m^3$ or $NO_2$ > 53ppb yearly has been found to be associated with increased risk of other respiratory illnesses.[83]

[h] All variables in this category are area-level variables at the level of the Census Dissemination Area.

[i] 1st quintile represents 0-2.1 persons/dwelling; 2nd quintile: 2.2 – 2.4 persons/dwelling; 3rd quintile: 2.5 – 2.6 persons/dwelling; 4th quintile: 2.7 – 3 persons/dwelling; 5th quintile: 3.1 – 5.7 persons/dwelling.

[j] 1st category represents 0 – 7.3% of buildings in the area being apartment buildings; 2nd category: 7.4 – 37.7% of buildings; 3rd category: 37.7 – 100% of buildings.

[k] 1st quintile represents 11.2 – 33.7% individuals uncoupled; 2nd quintile: 33.7 – 38.4% individuals; 3rd quintile: 38.5 – 43.6% individuals; 4th quintile: 43.6 – 51.0% individuals; 5th quintile: 51.0 – 94.6% individuals.

[l] 1st quintile represents 0 – 32.5% of working individuals in the area self-identifying as working in an essential job; 2nd quintile: 32.5 – 42.3% of individuals; 3rd quintile: 42.3 – 49.8% of individuals; 4th quintile: 50.0 – 57.5% of individuals; 5th quintile: 57.5 – 114.3% of individuals.

[m] Income quintile has variable cut-off values in each city or census area, to take cost of living into account. A DA being in quintile 1 means it is among the lowest 20% DAs in its city by income.

[n] 1st quintile represents 0 – 4.1% of individuals 25-64 years old without diploma; 2nd quintile: 4.1 – 7.5% of individuals; 3rd quintile: 7.5 – 11.4% of individuals; 4th quintile: 11.4 – 17.1% of individuals; 5th quintile: 17.1 – 94.3% of individuals.

[o] 1st quintile represents 0 – 2.2% of individuals in the area self-identifying as a visible minority; 2nd quintile: 2.2 – 7.5% of individuals; 3rd quintile: 7.5 – 18.7% of individuals; 4th quintile: 18.7 – 43.5% of individuals; 5th quintile: 43.5 – 100% of individuals.

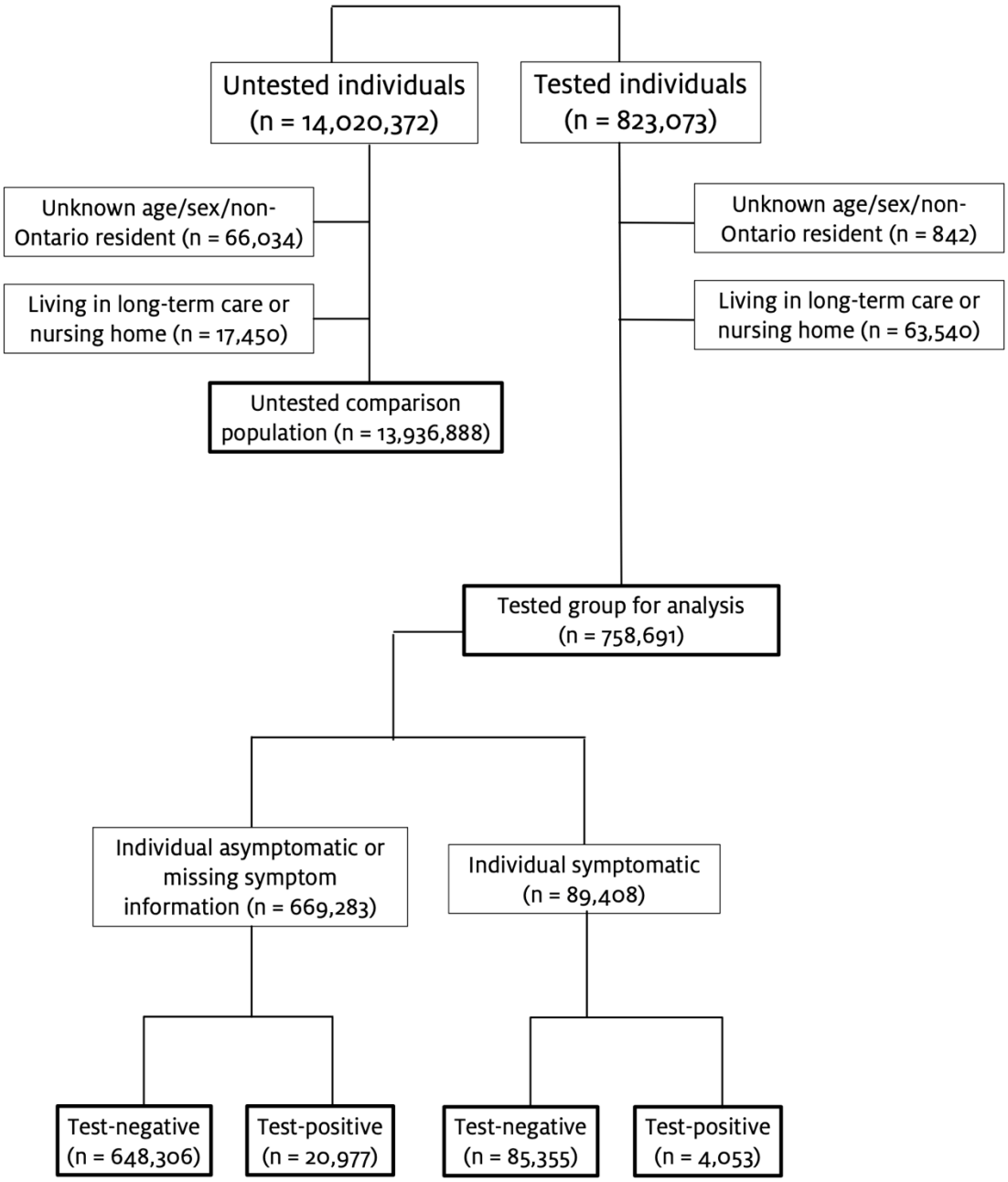
[p] 1st category represents 0 – 2.1% of individuals in DA being recent immigrants; 2nd category: 2.1 – 4.7% of individuals; 3rd category: 4.7 – 41.2% of individuals.



medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

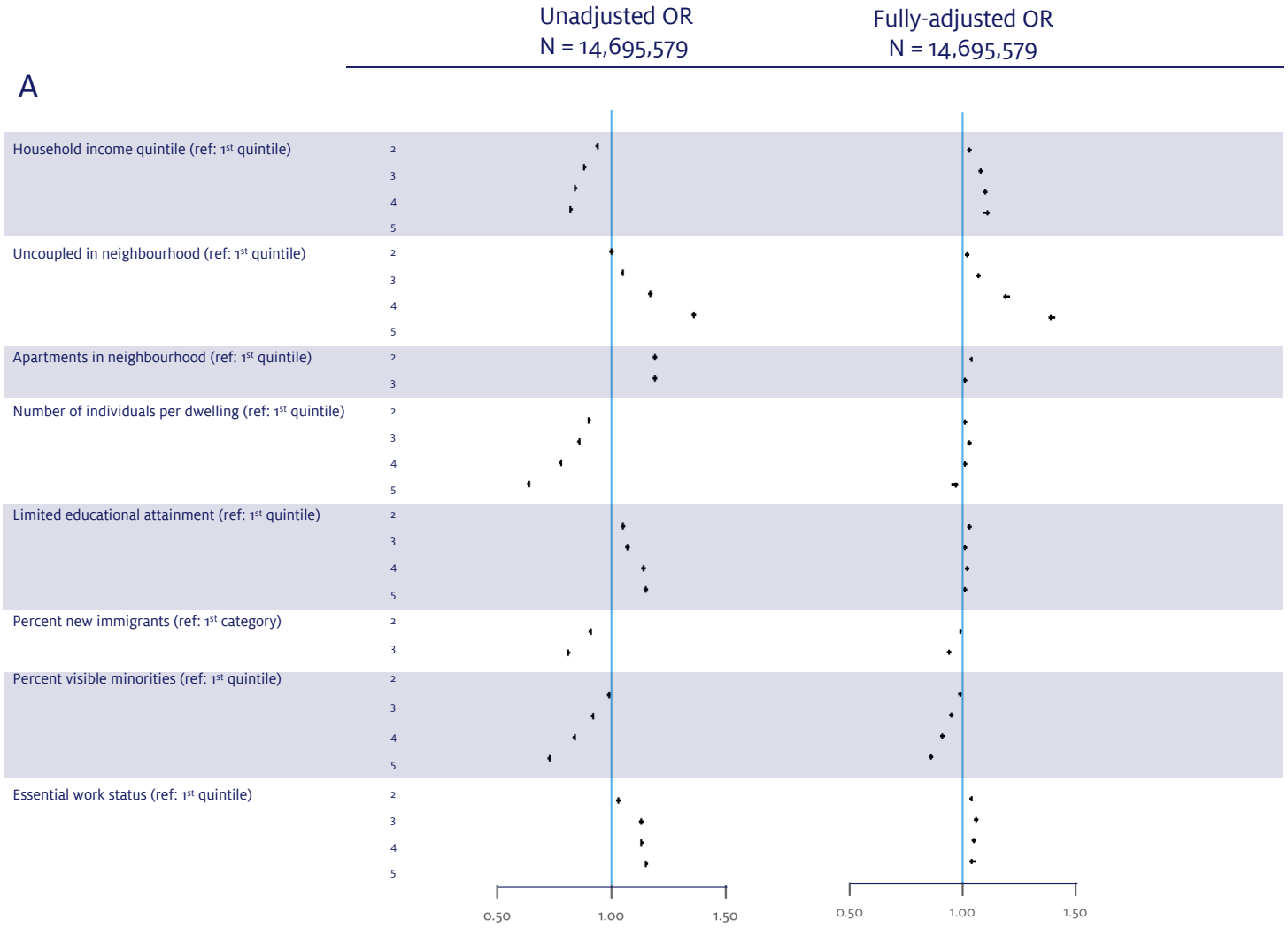
**Figure 1.** Diagram of inclusion and exclusion criteria and resulting analytic datasets.




Individuals were excluded if they had a listed age >105 years, had a listed postal code outside of Ontario, or had no record of contact with the healthcare system in the past 8 years (since March 2012).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

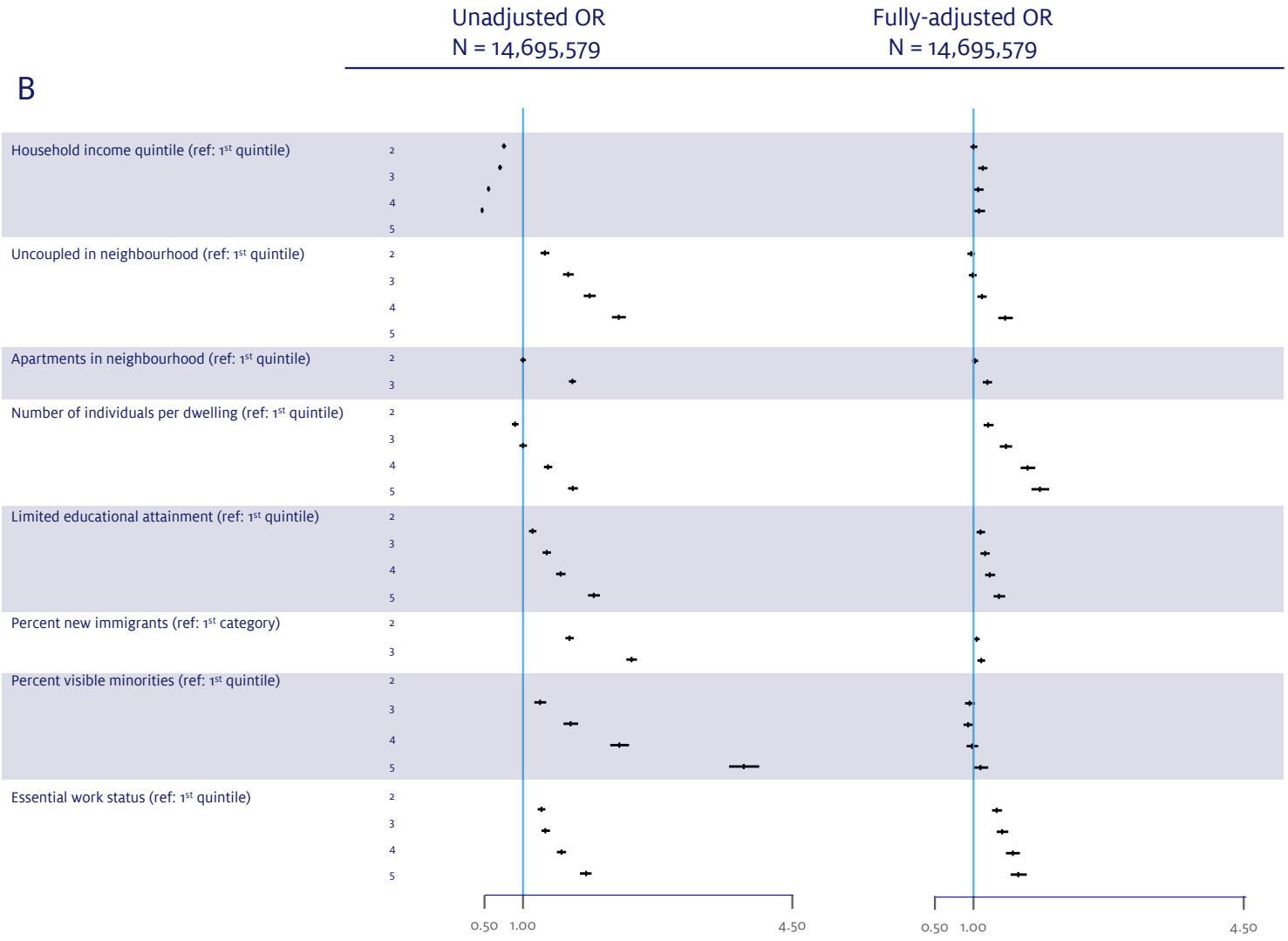
683 **Figure 2**. Unadjusted and fully-adjusted association between social determinants and SARS-CoV-2 testing (panel A) and
684 COVID-19 diagnosis (panel B), using the case-control design, in Ontario (March 1 to June 20, 2020).




Forest plots show the odds of being tested for SARS-CoV-2, compared to the reference group (1st quintile or 1st category). In all cases, the first quintile/category represents the lowest quintile of values for the variable at the dissemination area (DA)-level. For example, the lowest quintile of household income represents individuals living in the lowest 20% of DAs by income quintile. Fully-adjusted models control for all variables listed in Table 2, including (but not limited to) all variables shown in this figure.

685

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted March 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.




Forest plots show the odds of being tested for SARS-CoV-2, compared to the reference group (1st quintile or 1st category). In all cases, the first quintile/category represents the lowest quintile of values for the variable at the dissemination area (DA)-level. For example, the lowest quintile of DA-level income represents individuals living in the lowest 20% of DAs by income quintile. Fully-adjusted models are shown in the rightmost column of Table 2 and control for all variables listed in Table 2, including (but not limited to) all variables shown in this figure.

686