medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 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

Word count: 2989 (not counting abstract, figures, & tables)

Abstract word count: 347

Key Points word count: 101

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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 November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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.

Key Points

Question: What are the social determinants of health that contextualize individual-level risks for coronavirus disease 2019 (COVID-19), and how do selection biases affect our understanding of these risks?

Findings: In this province-wide observational study of 14.7 million Canadians, social determinants related to housing, education, and recent immigration were associated with increased COVID-19 risks, with little evidence of selection bias. Individual factors, such as underlying health conditions, were more prone to selection bias using certain analytic approaches.

Meaning: Social determinants of health appear to drive COVID-19 incidence in Ontario, Canada. Interventions aiming to prevent COVID-19 transmission should address these empiric structural risks.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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

Importance: 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.

Objective: Characterizing individual, environmental, and social determinants of SARS-CoV-2 testing and COVID-19 diagnosis.

Design: We conducted cross-sectional analyses among 14.7 million people comparing individual, environmental, and social determinants among individuals who were tested versus not yet tested. Among those diagnosed, we used three analytic designs to compare predictors of: 1) individuals testing positive versus negative; 2) symptomatic individuals testing positive versus 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.

Setting: Ontario, Canada.

Participants: All individuals with ≥1 healthcare system contact since March 2012, excluding individuals deceased before, or born after, March 1, 2020, or residing in a long-term care facility.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

Exposures: Individual-level characteristics (age, sex, underlying health conditions, prior healthcare use), area-based environmental (air pollution) exposures, and area-based social determinants of health (income, education, housing, marital status, race/ethnicity, and recent immigration).

Main Outcomes and Measures: Odds of SARS-CoV-2 test, and of COVID-19 diagnosis.

Results: Of a total of 14,695,579 individuals, 758,691 had been tested, 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 design. Residing in areas with highest household density (adjusted odds ratio: 2.08; 95%CI: 1.95-1.21), lowest educational attainment (adjusted odds ratio: 1.52; 95%CI: 1.44-1.60), and highest proportion of recent immigrants (adjusted odds ratio: 1.12; 95%CI: 1.07-1.16) were consistently related to increased odds of COVID-19 across analytic designs.

Conclusions and Relevance: 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 and systemic racism.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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] There has been greater focus on characterizing individual determinants of COVID-19, such as age,[2,5] sex,[6,7] and underlying health conditions,[2,8] than on social determinants of health.[9]

Heterogeneity in social determinants of COVID-19 may exist at the individual and network levels (for example, by housing density[10-12]). In addition, social determinants including barriers to healthcare, systemic racism, and xenophobia have been implicated in COVID-19 risk.[13,14] 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.[15] While previous studies suggest that air pollution may be biologically related to COVID-19,[16] including severe COVID-19,[17,18] it may also play a role in COVID-19 risk by operating within the context of low-quality housing and environmental racism.[19,20]

Using observational data to identify determinants of COVID-19 relies on SARS-CoV-2 testing, which is not equally distributed.[7] Differential testing introduces the potential for selection biases,[21,22] including collider bias.[22] Collider bias may be introduced into observational studies if the determinants under investigation are related to both COVID-19 and the likelihood of testing.[22-24]

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

Thus, to identify heterogeneity across potential determinants of COVID-19, we examined individual, environmental, and social determinants associated with SARS-CoV-2 testing and diagnoses among 14.7 million individuals in Ontario, Canada. We compared three analytic approaches to examine the role of selection biases.[22]

Methods

Study population, setting, and design

We conducted an observational study using population-based laboratory and health administrative databases from Ontario, Canada. Ontario has a single-payer health system that provides universal access to hospital and physician services[25] and laboratory testing.[26] We used data from individuals tested between March 1 and June 20, 2020 to identify determinants associated with testing, and then used three analytic designs to identify determinants associated with laboratory-confirmed diagnosis of COVID-19.

Data sources, linkages, and inclusion criteria

We linked individual-level SARS-CoV-2 testing data from the Ontario Laboratories Information System (OLIS) to relevant health-related datasets containing demographic, healthcare use, and area-level information (**eTable 1**). These datasets were linked using unique encoded identifiers and analyzed at ICES, a not-for-profit research institute in Ontario.[27]

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

OLIS captured approximately 88% of all laboratory-confirmed COVID-19 cases reported by the province during the study period. 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.[28] Further details are available below and in **eTable 1**.

To assess community transmission risks, 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 (**Figure 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, and were not residing in a long-term care facility. A diagram of individuals included and excluded is shown in **Figure 1**.

Selection and definition of potential determinants of COVID-19 diagnoses

Individual-level determinants included sex, age group, underlying health conditions, and prior healthcare use (**eTable 1**). Underlying health conditions were selected as: a)

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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 identified in the literature as associated with COVID-19 severity[2,29-32] or associated with symptoms similar to those caused by COVID-19, because severity and symptoms may lead to differential testing and thus, selection bias;[33-38] or b) health conditions such as dementia that increase the need for personal care support, thereby reflecting network-level contact patterns that intersect with occupational risks among essential care providers.[39,40]

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[41] Aggregated Diagnostic Groups (ADGs)[42] as a composite measure of comorbidities.

Environmental determinants included fine particulate matter ($PM_{2.5}$) using satellite-derived estimates[43] and land-use regression model for NO_2,[44] at the postal-code level.

Social determinants were conceptualized as area-based variables that may signal contact rates within and/or outside households [household density, uncoupled (for example, not married) status];[45,46] socio-economic barriers to healthcare access and/or housing or proxies of occupation (household income, educational attainment, apartment building density, and social assistance);[47,48] and unmeasured confounders related to race/ethnicity (visible minority status and recent immigration).[13,14] These determinants were derived from available measures within the long-form 2016 Canada Census at the

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

level of dissemination areas (DA), which are 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, the high frequency of zeros permitted the creation of only three categories (i.e., comparing the fourth and fifth quintiles with the lowest 60%). Variable definitions are detailed in **eTable1**.

Statistical analysis

To identify determinants of testing, the outcome was defined as receipt of at least one SARS-CoV-2 test during the study period. The comparator group comprised Ontario residents who did not have a 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. These are geographic areas in which public health measures were differentially applied[50] and along which there may be considerable variability in measured and unmeasured social determinants.[51] We opted to include public health regions *a priori* in the model due to these predicted variations in public health response.

To identify determinants of COVID-19, the outcome variable was a laboratory-confirmed diagnosis. To explore the potential for effects of a variety of sources of bias from testing, we compared the results according to different comparison sub-groups, by

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

characteristics that may have influenced the probability of being tested. We first compared individuals who tested positive to individuals who tested negative ("pseudo-test-negative design"). Second, we restricted the analysis to individuals with symptomatic illness, mirroring the test-negative design commonly used to assess influenza vaccine effectiveness (the "true test-negative design").[52] Third, we compared all individuals who tested positive to all individuals who did *not* test positive (i.e., to individuals who tested negative or were not tested; the "case-control design"). Potential mechanisms of selection biases in each analytic design are outlined in **eFigure 1A-C**.

We then conducted unadjusted, age/sex-adjusted, and fully-adjusted logistic regression models (including all determinants and public health region, as with testing) using each of the three analytic approaches to identify heterogeneity in individual, environmental, and social determinants of COVID-19 diagnosis. We interpreted each set of determinants (individual, environmental, social) as independent analyses based on the directed acyclic graphs (**eFigure 1A-C**). Statistical analyses were conducted using SAS v. 9.4 (Cary, NC).

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

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

Of 758,691 individuals tested during the study period, 25,030 (3.3%) received a laboratory-confirmed diagnosis (**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. **Table 1** describes the characteristics of the study population.

Determinants of SARS-CoV-2 testing

In fully-adjusted analyses, odds of testing increased with age (**Table 2**). Males had lower odds of testing than females. Nearly every underlying health condition was associated with increased odds of testing, as were most measures of prior healthcare use. In contrast, higher measures of ambient air pollution were associated with reduced odds of testing. There was little variability in the odds of testing by most area-based social determinants. 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 across nearly all determinants appeared to be progressively attenuated from unadjusted, to age/sex-adjusted, to fully-adjusted regression models (**eTable 2**).

Variability in determinants of COVID-19 diagnosis across three analytic approaches

Choice of analytic design and comparison group restrictions had substantial influences on the magnitude and direction of the association between individual-level determinants and diagnosis (**Table 2**). Determinants associated with odds of testing further from the null appeared to be more variable in terms of odds of diagnosis across analytic designs.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

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]=5.57; 95%CI, 5.45-5.70), and the odds of diagnosis differed between the pseudo-test-negative design (aOR=1.72; 95%CI, 1.48-2.01), the true test-negative design (aOR=5.95; 95%CI, 3.57-9.92), and the case-control design (aOR=7.09; 95%CI, 6.08-8.26). Some health conditions associated with higher odds of testing, such as chronic respiratory conditions, and indicators of prior healthcare use appeared protective against COVID-19 using the pseudo-test-negative design, but reverted to the null or showed increased odds of diagnosis using the case-control design. As with testing, effect measures for COVID-19 diagnosis across all determinants appeared to be progressively attenuated from unadjusted, to age/sex-adjusted, to fully-adjusted regression models (**eTables 3-5**).

Individual determinants of COVID-19 diagnosis using the case-control design

Using the case-control design, age, certain comorbidities (e.g., hypertension, diabetes, dementia), and increased prior healthcare use were associated with increased odds of COVID-19. In contrast, other comorbidities (e.g., cancer, substance abuse) and receipt of influenza vaccine in the 2019-2020 season were associated with reduced odds of diagnosis (**Table 2**).

Environmental determinants of COVID-19 diagnosis using the case-control design

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 consistently across analytic approaches.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

Social determinants of COVID-19 diagnosis using the case-control design

Area-level social determinants independently associated with COVID-19 included higher household income, increased receipt of social assistance, lower educational attainment, greater percentages of uncoupled individuals, higher household density, increased apartment building density, and greater percentages of recent immigrants.

Figures 2A and **2B** highlight the changes from unadjusted to fully-adjusted models for the associations between area-level social determinants and testing and diagnosis, respectively. For testing, the associations were attenuated after adjustment for nearly all determinants. The association for household income reversed direction following adjustment. For diagnosis, the associations were attenuated after adjustment for all social determinants except for household density. Similar to testing, the association with COVID-19 for household income reversed direction following adjustment.

Discussion

In this large-scale, population-based study, we identified social determinants as key potential determinants for a diagnosis of COVID-19. We identified variability in associations, likely due to selection biases, by analytic design across individual determinants of COVID-19 diagnosis.

We identified increased odds of diagnosis associated with household density, educational attainment, uncoupled status, and recent immigration, consistent with

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

findings from other settings.[51,53,54] In particular, the results concerning social determinants and COVID-19 risk presented here suggest these social determinants are not likely artifacts of selection bias from study design, given the consistency of findings across different analytic approaches. Lower educational attainment or uncoupled status may be associated with higher exposure risk through lower-paying jobs in the service industry[55] and/or other high-exposure occupations, either because those jobs cannot be done feasibly with proper protections, or because protective policies fail to be issued, leaving workers at high risk.[56,57] Higher percentages of recent immigrants in an area were associated with COVID-19 diagnosis, even after adjustment, although the percentage of visible minorities was not. Both variables might represent residual measures of mediators of systemic racism, potentiating relative risks of SARS-CoV-2 exposure and COVID-19 severity,[58-60] including COVID-19-related hospitalization/death.[14,21,30,54] We found that an association between visible minority and COVID-19 diagnosis in Ontario was attenuated after adjustment for individual, environmental, and other social determinants. These findings likely reflect what is already known about race and ethnicity as social constructs and social determinants of health.[61] This is further supported by the attenuation of effect estimates from the unadjusted, to age/sex-adjusted, to fully-adjusted models. Finally, the fact that there was little variation in the odds of testing rates across levels of most of the social determinants suggests that testing resources may not be prioritizing those who are most at risk.[62] However, even with this finding, our results indicate that social determinants of health play important roles in determining COVID-19 risk. Taken together, our findings suggest a need to increase and/or redirect resources that specifically address social

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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 such as household density[48,63] (e.g., voluntary temporary isolation hotels[64]), proxies for occupational risk[56,60] (e.g., paid sick leave[65]), and other mediators of systemic racism[62,66,67] (e.g. community-led outreach testing[68]).

We also identified the size and direction of influence that selection biases may have, adding to the ongoing conversation on the challenges of interpreting determinants for COVID-19 due to collider bias.[22] Although we did not directly explore each potential collider mechanism, the pattern that emerged was that underlying health conditions associated with COVID-19 severity[2] may have been prone to collider bias, as evidenced by higher testing and a seemingly protective effect when using the test-negative designs that reversed when using the case-control design. It is, however, also possible that a high number of covariates in fully-adjusted models caused instability in effect measure estimates for these covariates. Age also demonstrated a similar potential susceptibility to collider bias, possibly mediated by COVID-19 severity and symptoms (i.e., older age groups are more likely to develop severe/symptomatic infection if infected[2]). 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.[33-36,38] 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 symptomatic and asymptomatic individuals, but that may have been partly due to the extensive (74.6%) missingness of symptom reporting in the testing data. We opted not to use multiple imputation methods to determine symptom information because, at the time the analysis was conducted,

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

reported symptoms of COVID-19 were highly variable and the extent to which individuals may have asymptomatic or pre-symptomatic illness was unclear, limiting our confidence in the generalizability of existing information to individuals whose information was missing. Thus, while there is an intuitive desire to compare test positivity rates among those tested—a common metric of comparison across surveillance reports and dashboards[69,70]—the risks of erroneously identifying a "risk/protective factor", particularly individual determinants such as underlying health conditions, sex, and age, are high and deserve careful interpretation by examining, as best possible, the reasons for testing.[22] In the context of low overall levels of testing, the case-control design may have mitigated some potential sources of selection bias, with the assumption that those not tested are similar to those who tested negative.[21,22] Additionally, we found that some underlying health conditions and prior healthcare use remained associated with diagnoses, reflecting either unmeasured confounders or possible biological susceptibility to infection if exposed;[15,16,71-73] and suggesting that strategies tailored to reduce exposures among individuals characterized by these individual determinants could also be important in risk-tailored prevention.

Finally, these analyses identified $PM_{2.5}$ as being related to the odds of COVID-19. It is likely that some of this effect is due to underlying social determinants of health and access to tests. Air pollution is often more intense, and results in worse health outcomes, in areas with higher social deprivation.[74,75] However, existing studies have also implicated environmental pollution as having a biological relationship to the risk and severity of COVID-19 and other respiratory infections.[15,16,19] Taken together with other

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

findings regarding social determinants of health, these results indicate the importance of considering air pollution and particulate matter in both individual- and area-directed COVID-19 interventions.

Limitations

Our diagnoses were restricted to laboratory-confirmed cases and to the 88% of diagnoses available via OLIS, and thus could miss probable cases as well as the remaining laboratory-confirmed cases that may have different 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.[76] Future analyses include examining changes in the direction and magnitude of determinants over the course of the outbreak. Although we generated directed acyclic graphs for the general categories of determinants to help conceptualize and mitigate selection bias while also identifying plausible determinants, it is possible that we over-adjusted when interpreting individual variables within each category.[77] 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 COVID-19 risk 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. Some important determinants identified in the prior literature, such as obesity,[8,78] were

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

not available for our study.[79] Others, such as occupation, were not included due to uncertainty related to quantifying and appropriately grouping the occupational classifications available in the Census data at the area level.

Conclusion

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

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

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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 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.

Acknowledgements: We thank IMS Brogan Inc. for use of their Drug Information Database. 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 ACK 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 November 12, 2020. 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.

440

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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. CDC COVID-19 Response Team. Coronavirus Disease 2019 in Children — United States, February 12-April 2, 2020. *Morbidity and Mortality Weekly Report.* 2020;69(14):422-426.
6. Takahashi T, Ellingson MK, Wong P, et al. Sex differences in immune responses that underlie COVID-19 disease outcomes. *Nature.* 2020.
7. 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.
8. 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.
9. Abrams E, Szefler S. COVID-19 and the impact of social determinants of health. *The New England journal of medicine.* 2020;382(21):2061-2063.
10. 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.*
11. 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.
12. Wang L, Mishra S. *The distribution of cumulative COVID-19 cases in Ontario by acquisition type over time.* 2020.
13. 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.
14. 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.
15. 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

16. 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.
17. 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.
18. 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.
19. 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.
20. Environmental racism: time to tackle social injustice. *The Lancet Planetary Health.* 2018;2(11).
21. 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.
22. Griffith G, Morris T, Tudball M, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *medRxiv.* 2020.
23. Hernán M, Robins J. *Causal Inference.* 2014.
24. Elwert F, Winship C. Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. *Annu Rev Sociol.* 2014;40:31-53.
25. Health care in Ontario: OHIP. https://www.ontario.ca/page/health-care-ontario. Published 2020. Accessed October 29, 2020.
26. *Schedule of benefits for laboratory services.* Ontario: Ministry of Health, Ontario Health Insurance Plan, Laboratories and Genetics Branch; July 1 2020.
27. IC/ES: Data, Discovery, Better Health. https://www.ices.on.ca. Published 2020. Accessed October 29, 2020.
28. 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.
29. Pennington E. Asthma increases risk of severity of COVID-19. *Cleve Clin J Med.* 2020.
30. Williamson EJ, Walker AJ, Bhaskaran K, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature.* 2020.
31. 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.
32. Merkler AE, Parikh NS, Mir S, et al. Risk of ischemic stroke in patients with COVID-19 versus patients with influenza. *medRxiv.* 2020.
33. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: May 14, 2020.* May 14 2020.
34. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: May 2, 2020.* May 2 2020.
35. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: April 15, 2020.* April 15 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

36. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: May 28, 2020.* May 28 2020.
37. Ontario Ministry of Health. *COVID-19 Quick Reference Public Health Guidance on Testing and Clearance: March 27, 2020.* 2020.
38. Ontario Ministry of Health. *COVID-19 Provincial Testing Guidance Update: April 8, 2020.* April 8, 2020 2020.
39. Tolea MI, Morris JC, Galvin JE. Trajectory of Mobility Decline by Type of Dementia. *Alzheimer Dis Assoc Disord.* 2016;30(1):60-66.
40. 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.
41. Johns Hopkins ACG(c) System. Johns Hopkins. https://www.hopkinsacg.org. Published 2020. Accessed November 1, 2020.
42. 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.
43. 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.
44. 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.
45. 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.
46. Park SY, Kim YM, Yi S, et al. Coronavirus Disease Outbreak in Call Center, South Korea. *Emerg Infect Dis.* 2020;26(8):1666-1670.
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. Lin SL. Intersectionality and inequalities in medical risk for severe COVID-19 in the Canadian Longitudinal Study on Aging. *Gerontologist.* 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

54. 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.
55. *Workplace Safety & Prevention Services Guidance on Health and Safety for the Personal Care Service Sector during COVID-19.* Ontario: Workplace Safety & Prevention Services;2020.
56. 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.
57. Contrera J. The N95 shortage America can't seem to fix. *Washington Post.* September 21, 2020, 2020.
58. 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.
59. 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.
60. 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).
61. Ordóñez CE. Not race, racism: Concerns of COVID-19 affecting African Americans. *Anthropological Responses to Heatlh Emergencies.* 2020.
62. Dodds C, Fakoya I. Covid-19: ensuring equality of access to testing for ethnic minorities. *BMJ.* 2020;369:m2122.
63. 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.
64. 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.
65. 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.
66. Webb Hooper M, Napoles AM, Perez-Stable EJ. COVID-19 and Racial/Ethnic Disparities. *JAMA.* 2020.
67. 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.
68. 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.
69. ICES COVID-19 Dashboard. ICES. https://www.ices.on.ca/DAS/AHRQ/COVID-19-Dashboard. Published 2020. Accessed October 29, 2020.
70. Ontario COVID-19 Data Tool. Public Health Ontario. https://www.publichealthontario.ca/en/data-and-analysis/infectious-disease/covid-

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

19-data-surveillance/covid-19-data-tool. Published 2020. Accessed October 29, 2020.

71. 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.
72. Hussain A, Bhowmik B, do Vale Moreira NC. COVID-19 and diabetes: Knowledge in progress. *Diabetes Res Clin Pract.* 2020;162:108142.
73. Elliot AJ, Harcourt SE, Hughes HE, et al. The COVID-19 pandemic: a new challenge for syndromic surveillance. *Epidemiol Infect.* 2020;148:e122.
74. 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.
75. Næss Ø, Piro FN, Nafstad P, Smith GD, Leyland AH. Air Pollution, Social Deprivation, and Mortality. *Epidemiology.* 2007;18(6):686-694.
76. 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.
77. Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *American journal of epidemiology.* 2013;177(4):292-298.
78. Kwok S, Adam S, Ho JH, et al. Obesity: A critical risk factor in the COVID-19 pandemic. *Clin Obes.* 2020:e12403.
79. ICES Data Dictionary. ICES. https://datadictionary.ices.on.ca/Applications/DataDictionary/Default.aspx. Published 2020. Accessed November 1, 2020.
80. In: NAAQS Table, ed. Washington, DC: United States Environmental Protection Agency.

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		Symptomatic tested individuals only		All individuals	
	Tested negative (N = 733,661)	**Tested positive (N = 25,030)**	**Tested negative and symptomatic (N = 85,355)**	**Tested positive and symptomatic (N = 4,053)**	**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	47.91 ± 20.80	50.83 ± 20.39	40.31 ± 22.91	49.45 ± 20.95
Age groups (years)						
0-4	10,870 (1.5%)	223 (0.9%)	1,627 (1.9%)	17 (0.4%)	727,305 (5.2%)	11,093 (1.5%)
5-19	33,642 (4.6%)	1,149 (4.6%)	4,017 (4.7%)	125 (3.1%)	2,395,518 (17.2%)	34,791 (4.6%)
20-34	157,086 (21.4%)	6,015 (24.0%)	19,431 (22.8%)	864 (21.3%)	2,810,230 (20.2%)	163,101 (21.5%)
35-49	160,506 (21.9%)	6,084 (24.3%)	20,495 (24.0%)	934 (23.0%)	2,772,204 (19.9%)	166,590 (22.0%)
50-64	192,178 (26.2%)	7,009 (28.0%)	21,329 (25.0%)	1,159 (28.6%)	2,865,926 (20.6%)	199,187 (26.3%)
65-74	84,077 (11.5%)	2,072 (8.3%)	8,765 (10.3%)	386 (9.5%)	1,394,043 (10.0%)	86,149 (11.4%)
75-84	52,655 (7.2%)	1,284 (5.1%)	5,181 (6.1%)	251 (6.2%)	713,639 (5.1%)	53,939 (7.1%)
≥85	42,647 (5.8%)	1,194 (4.8%)	4,510 (5.3%)	317 (7.8%)	258,023 (1.9%)	43,841 (5.8%)
Male sex: N (%)	289,769 (39.5%)	11,627 (46.5%)	33,078 (38.8%)	1,898 (46.8%)	6,930,233 (49.7%)	301,396 (39.7%)
Living in rural/small town: N (%)[a]	82,034 (11.2%)	887 (3.5%)	14,140 (16.6%)	304 (7.5%)	1,390,426 (10.0%)	82,921 (10.9%)
Public Health region						
Central East	59,951 (8.2%)	707 (2.8%)	7,898 (9.3%)	193 (4.8%)	874,010 (6.3%)	60,658 (8.0%)
Central West	136,116 (18.6%)	3,375 (13.5%)	17,469 (20.5%)	951 (23.5%)	2,683,999 (19.3%)	139,491 (18.4%)
Durham	30,618 (4.2%)	1,244 (5.0%)	1,341 (1.6%)	125 (3.1%)	671,605 (4.8%)	31,862 (4.2%)
Eastern	53,097 (7.2%)	466 (1.9%)	9,752 (11.4%)	152 (3.8%)	829,504 (6.0%)	53,563 (7.1%)
North	56,977 (7.8%)	308 (1.2%)	14,442 (16.9%)	177 (4.4%)	761,495 (5.5%)	57,285 (7.6%)
Ottawa	52,754 (7.2%)	1,360 (5.4%)	3,761 (4.4%)	190 (4.7%)	980,501 (7.0%)	54,114 (7.1%)
Peel	63,093 (8.6%)	4,464 (17.8%)	5,780 (6.8%)	376 (9.3%)	1,465,044 (10.5%)	67,557 (8.9%)
South West	87,412 (11.9%)	2,040 (8.2%)	9,101 (10.7%)	673 (16.6%)	1,609,925 (11.6%)	89,452 (11.8%)
Toronto	140,217 (19.1%)	8,728 (34.9%)	10,360 (12.1%)	797 (19.7%)	2,761,786 (19.8%)	148,945 (19.6%)
York	48,565 (6.6%)	2,119 (8.5%)	5,068 (5.9%)	376 (9.3%)	1,157,124 (8.3%)	50,684 (6.7%)
Underlying chronic health conditions						
Asthma	136,556 (18.6%)	3,887 (15.5%)	17,831 (20.9%)	671 (16.6%)	2,085,118 (15.0%)	140,443 (18.5%)
COPD	34,261 (4.7%)	715 (2.9%)	4,616 (5.4%)	168 (4.1%)	249,636 (1.8%)	34,976 (4.6%)

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

Hypertension	226,111 (30.8%)	7,488 (29.9%)	24,364 (28.5%)	1,313 (32.4%)	2,836,790 (20.4%)	233,599 (30.8%)
Diabetes	108,758 (14.8%)	4,268 (17.1%)	11,763 (13.8%)	684 (16.9%)	1,363,667 (9.8%)	113,026 (14.9%)
Congestive heart failure	35,947 (4.9%)	983 (3.9%)	3,864 (4.5%)	194 (4.8%)	225,153 (1.6%)	36,930 (4.9%)
Dementia or frailty	31,804 (4.3%)	1,106 (4.4%)	3,399 (4.0%)	296 (7.3%)	124,295 (0.9%)	32,910 (4.3%)
Cancer[b]	24,333 (3.3%)	484 (1.9%)	2,333 (2.7%)	89 (2.2%)	213,036 (1.5%)	24,817 (3.3%)
Chronic kidney disease[b]	37,630 (5.1%)	1,122 (4.5%)	3,581 (4.2%)	199 (4.9%)	254,541 (1.8%)	38,752 (5.1%)
Immunocompromised[c]	13,325 (1.8%)	319 (1.3%)	1,496 (1.8%)	57 (1.4%)	107,824 (0.8%)	13,644 (1.8%)
Advanced liver disease	9,020 (1.2%)	212 (0.8%)	1,045 (1.2%)	35 (0.9%)	78,747 (0.6%)	9,232 (1.2%)
Ischemic heart disease	33,855 (4.6%)	880 (3.5%)	3,773 (4.4%)	168 (4.1%)	324,682 (2.3%)	34,735 (4.6%)
Ischemic stroke or TIA[d]	13,777 (1.9%)	411 (1.6%)	1,531 (1.8%)	77 (1.9%)	99,503 (0.7%)	14,188 (1.9%)
Schizophrenia[e]	7,526 (1.0%)	220 (0.9%)	732 (0.9%)	42 (1.0%)	64,494 (0.5%)	7,746 (1.0%)
Substance abuse[e]	22,238 (3.0%)	478 (1.9%)	3,001 (3.5%)	59 (1.5%)	216,156 (1.6%)	22,716 (3.0%)
Healthcare use						
Adjusted Diagnostic Group quintile						
1 (0 ADGs)	31,959 (4.4%)	1,571 (6.3%)	3,006 (3.5%)	247 (6.1%)	1,616,852 (11.6%)	33,530 (4.4%)
2 (1-2 ADGs)	101,895 (13.9%)	3,601 (14.4%)	11,419 (13.4%)	595 (14.7%)	3,150,969 (22.6%)	105,496 (13.9%)
3 (3-4 ADGs)	149,173 (20.3%)	5,101 (20.4%)	17,220 (20.2%)	782 (19.3%)	3,385,602 (24.3%)	154,274 (20.3%)
4 (5-6 ADGs)	150,394 (20.5%)	5,115 (20.4%)	17,736 (20.8%)	822 (20.3%)	2,626,403 (18.8%)	155,509 (20.5%)
5 (7-27 ADGs)	300,240 (40.9%)	9,642 (38.5%)	35,974 (42.1%)	1,607 (39.6%)	3,157,062 (22.7%)	309,882 (40.8%)
Hospital admissions, past 3 years: Mean (SD)	0.44 ± 1.34	0.33 ± 1.21	0.45 ± 1.29	0.36 ± 1.09	0.19 ± 0.60	0.44 ± 1.34
Outpatient physician visits, past year: Mean (SD)	7.35 ± 8.68	7.06 ± 8.33	7.15 ± 8.40	7.07 ± 8.64	4.67 ± 6.05	7.34 ± 8.67
Influenza vaccination (2019-2020 season)	213,722 (29.1%)	5,547 (22.2%)	23,429 (27.4%)	1,019 (25.1%)	2,978,472 (21.4%)	219,269 (28.9%)
Environmental determinants[f]						
$PM_{2.5}$ category (μg/m^3 yearly)						
2 to <6	161,300 (22.0%)	1,831 (7.3%)	27,764 (32.5%)	556 (13.7%)	2,481,201 (17.8%)	163,131 (21.5%)
6 to <7	91,134 (12.4%)	1,766 (7.1%)	10,930 (12.8%)	318 (7.8%)	1,555,790 (11.2%)	92,900 (12.2%)
7 to <8	207,966 (28.3%)	8,476 (33.9%)	21,183 (24.8%)	1,281 (31.6%)	4,450,218 (31.9%)	216,442 (28.5%)
8 to <9	211,861 (28.9%)	11,127 (44.5%)	19,639 (23.0%)	1,363 (33.6%)	4,258,069 (30.6%)	222,988 (29.4%)
≥9	59,674 (8.1%)	1,747 (7.0%)	5,749 (6.7%)	517 (12.8%)	1,073,665 (7.7%)	61,421 (8.1%)
NO_2 category (parts per billion yearly)						
0 to 6	328,613 (44.8%)	5,237 (20.9%)	48,579 (56.9%)	1,500 (37.0%)	5,505,976 (39.5%)	333,850 (44.0%)

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

6 to 8	170,693 (23.3%)	5,599 (22.4%)	17,477 (20.5%)	1,079 (26.6%)	3,409,506 (24.5%)	176,292 (23.2%)
≥8	232,629 (31.7%)	14,111 (56.4%)	19,209 (22.5%)	1,456 (35.9%)	4,903,460 (35.2%)	246,740 (32.5%)
Social determinants of health[g] (area-level)						
Household income quintile[h]						
1 (lowest income)	156,320 (21.3%)	7,000 (28.0%)	17,159 (20.1%)	857 (21.1%)	2,679,780 (19.2%)	163,320 (21.5%)
2	148,687 (20.3%)	5,288 (21.1%)	17,057 (20.0%)	757 (18.7%)	2,698,807 (19.4%)	153,975 (20.3%)
3	145,317 (19.8%)	5,084 (20.3%)	16,771 (19.6%)	756 (18.7%)	2,791,340 (20.0%)	150,401 (19.8%)
4	140,352 (19.1%)	4,019 (16.1%)	17,162 (20.1%)	730 (18.0%)	2,809,529 (20.2%)	144,371 (19.0%)
5 (highest income)	138,103 (18.8%)	3,419 (13.7%)	16,822 (19.7%)	910 (22.5%)	2,815,238 (20.2%)	141,522 (18.7%)
Social assistance quintile (percentage of individuals in the area receiving governmental transfer payments)						
1 (38.0 - 63.5%)	141,146 (19.2%)	3,829 (15.3%)	15,784 (18.5%)	764 (18.9%)	3,016,045 (21.6%)	144,975 (19.1%)
2 (63.5 - 67.2%)	145,846 (19.9%)	4,719 (18.9%)	16,188 (19.0%)	794 (19.6%)	3,004,958 (21.6%)	150,565 (19.8%)
3 (67.2 - 70.8%)	137,208 (18.7%)	4,987 (19.9%)	16,155 (18.9%)	851 (21.0%)	2,658,737 (19.1%)	142,195 (18.7%)
4 (70.8 - 75.3%)	136,423 (18.6%)	4,952 (19.8%)	16,691 (19.6%)	736 (18.2%)	2,433,304 (17.5%)	141,375 (18.6%)
5 (75.3 - 98.1%)	160,519 (21.9%)	6,092 (24.3%)	19,144 (22.4%)	822 (20.3%)	2,571,707 (18.5%)	166,611 (22.0%)
Limited educational attainment quintile (percentage of adults aged 25-64 in the area not having received any diploma)						
1 (0.0 - 4.1%)	144,457 (19.7%)	3,903 (15.6%)	15,245 (17.9%)	709 (17.5%)	2,926,432 (21.0%)	148,360 (19.6%)
2 (4.1 - 7.5%)	154,215 (21.0%)	4,477 (17.9%)	17,690 (20.7%)	869 (21.4%)	2,978,444 (21.4%)	158,692 (20.9%)
3 (7.5 - 11.4%)	151,457 (20.6%)	5,052 (20.2%)	17,765 (20.8%)	759 (18.7%)	2,888,525 (20.7%)	156,509 (20.6%)
4 (11.4 - 17.1%)	148,159 (20.2%)	5,314 (21.2%)	18,014 (21.1%)	847 (20.9%)	2,659,544 (19.1%)	153,473 (20.2%)
5 (17.1 - 94.3%)	129,580 (17.7%)	6,010 (24.0%)	16,160 (18.9%)	824 (20.3%)	2,333,134 (16.7%)	135,590 (17.9%)
Uncoupled quintile (percent of individuals)						
1 (11.2 - 33.7%)	150,044 (20.5%)	3,539 (14.1%)	20,243 (23.7%)	824 (20.3%)	3,130,907 (22.5%)	153,583 (20.2%)
2 (33.7 - 38.4%)	128,561 (17.5%)	3,915 (15.6%)	15,262 (17.9%)	779 (19.2%)	2,698,847 (19.4%)	132,476 (17.5%)
3 (38.5 - 43.6%)	127,689 (17.4%)	4,632 (18.5%)	14,539 (17.0%)	693 (17.1%)	2,579,005 (18.5%)	132,321 (17.4%)
4 (43.6 - 51.0%)	145,560 (19.8%)	5,591 (22.3%)	15,894 (18.6%)	713 (17.6%)	2,633,710 (18.9%)	151,151 (19.9%)
5 (51.0 - 94.6%)	174,985 (23.9%)	7,037 (28.1%)	18,781 (22.0%)	976 (24.1%)	2,732,838 (19.6%)	182,022 (24.0%)

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

Household density quintile (range of persons per dwelling)						
1 (0 – 2.1)	162,623 (22.2%)	3,639 (14.5%)	20,066 (23.5%)	672 (16.6%)	2,474,391 (17.8%)	166,262 (21.9%)
2 (2.2 - 2.4)	140,653 (19.2%)	3,104 (12.4%)	19,406 (22.7%)	662 (16.3%)	2,368,013 (17.0%)	143,757 (18.9%)
3 (2.5 - 2.6)	104,896 (14.3%)	2,721 (10.9%)	13,059 (15.3%)	531 (13.1%)	1,866,317 (13.4%)	107,617 (14.2%)
4 (2.7 - 3)	166,089 (22.6%)	6,321 (25.3%)	17,843 (20.9%)	1,096 (27.0%)	3,291,097 (23.6%)	172,410 (22.7%)
5 (3.1 - 5.7)	152,578 (20.8%)	8,929 (35.7%)	14,345 (16.8%)	1,024 (25.3%)	3,775,489 (27.1%)	161,507 (21.3%)
Apartment building density category (percent of buildings that are apartment buildings in the area[i,j]						
1 (0 – 7.3%)	391,477 (53.4%)	12,377 (49.4%)	46,693 (54.7%)	2,243 (55.3%)	7,994,323 (57.4%)	403,854 (53.2%)
2 (7.4 - 37.7%)	145,108 (19.8%)	3,874 (15.5%)	19,348 (22.7%)	739 (18.2%)	2,478,855 (17.8%)	148,982 (19.6%)
3 (37.7 - 104%)	190,244 (25.9%)	8,463 (33.8%)	18,678 (21.9%)	1,003 (24.7%)	3,301,839 (23.7%)	198,707 (26.2%)
Recent immigration category (percentage of individuals in the area who are recent immigrants[k]						
1 (0.0 - 2.1%)	401,300 (54.7%)	8,271 (33.0%)	54,845 (64.3%)	2,002 (49.4%)	6,967,468 (50.0%)	409,571 (54.0%)
2 (2.1 - 4.7%	146,772 (20.0%)	5,409 (21.6%)	14,834 (17.4%)	852 (21.0%)	2,858,064 (20.5%)	152,181 (20.1%)
3 (4.7 - 41.2%)	174,829 (23.8%)	11,018 (44.0%)	14,393 (16.9%)	1,133 (28.0%)	3,898,463 (28.0%)	185,847 (24.5%)
Visible minority quintile (percentage of individuals in the area self-identifying as a visible minority)						
1 (0.0 - 2.2%)	130,912 (17.8%)	1,716 (6.9%)	21,800 (25.5%)	555 (13.7%)	2,115,641 (15.2%)	132,628 (17.5%)
2 (2.2 - 7.5%)	137,826 (18.8%)	2,233 (8.9%)	19,869 (23.3%)	677 (16.7%)	2,255,245 (16.2%)	140,059 (18.5%)
3 (7.5 - 18.7%)	137,744 (18.8%)	3,201 (12.8%)	16,345 (19.1%)	792 (19.5%)	2,451,335 (17.6%)	140,945 (18.6%)
4 (18.7 - 43.5%)	153,503 (20.9%)	5,466 (21.8%)	13,581 (15.9%)	887 (21.9%)	3,023,752 (21.7%)	158,969 (21.0%)
5 (43.5 - 102%)[l]	167,893 (22.9%)	12,140 (48.5%)	13,280 (15.6%)	1,097 (27.1%)	3,940,245 (28.3%)	180,033 (23.7%)

[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.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

[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.[80] $PM_{2.5}$ 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] 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.
[i] The census counts for apartment buildings are randomly rounded up or down to the nearest number divisible by 5, which causes some minor imprecision.
[j] 7.3% represents the 60th percentile.
[k] 2.1% represents the 60th percentile.
[l] Census counts for individuals are randomly rounded up or down to the nearest number divisible by 5, which causes some minor imprecision.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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 three analytic approaches, in Ontario, Canada between March 1 and June 20, 2020.

	Tested for SARS-CoV-2	Diagnosed with COVID-19		
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. test-negative among symptomatic individuals only (True 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)[a]**
Sample size	14,695,579	758,691	89,408	14,695,579
Individual determinants				
Age groups (years) (ref: 0 – 4)				
5-19	1.15 (1.12, 1.18)	1.92 (1.66, 2.23)	2.93 (1.75, 4.91)	1.72 (1.48, 1.99)
20-34	4.13 (4.05, 4.22)	2.00 (1.75, 2.30)	4.02 (2.47, 6.55)	7.01 (6.12, 8.04)
35-49	4.13 (4.05, 4.21)	1.98 (1.72, 2.27)	4.20 (2.58, 6.84)	6.73 (5.87, 7.71)
50-64	4.20 (4.12, 4.29)	2.01 (1.75, 2.31)	5.46 (3.35, 8.90)	6.98 (6.08, 8.00)
65-74	3.03 (2.97, 3.10)	1.60 (1.38, 1.85)	5.05 (3.07, 8.31)	3.94 (3.41, 4.55)
75-84	3.00 (2.94, 3.07)	1.61 (1.38, 1.87)	5.37 (3.23, 8.93)	3.86 (3.33, 4.48)
≥85	5.57 (5.45, 5.70)	1.72 (1.48, 2.01)	5.95 (3.57, 9.92)	7.09 (6.08, 8.26)
Male sex	0.76 (0.76, 0.76)	1.26 (1.23, 1.30)	1.35 (1.26, 1.44)	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 (0.65, 0.88)	0.76 (0.70, 0.83)
Public Health region (ref: Toronto)				
Central East	1.12 (1.10, 1.14)	0.40 (0.36, 0.45)	0.54 (0.42, 0.71)	0.44 (0.39, 0.50)
Central West	0.97 (0.96, 0.98)	0.61 (0.57, 0.65)	0.91 (0.78, 1.07)	0.58 (0.55, 0.62)
Durham	0.87 (0.86, 0.88)	0.95 (0.88, 1.02)	1.55 (1.24, 1.94)	0.81 (0.76, 0.87)
Eastern	1.06 (1.05, 1.08)	0.33 (0.29, 0.38)	0.39 (0.30, 0.51)	0.35 (0.31, 0.39)
North	1.19 (1.17, 1.21)	0.19 (0.16, 0.22)	0.30 (0.23, 0.39)	0.23 (0.19, 0.26)
Ottawa	0.97 (0.95, 0.99)	0.82 (0.75, 0.91)	1.06 (0.82, 1.37)	0.78 (0.70, 0.86)
Peel	0.98 (0.97, 0.99)	0.99 (0.95, 1.04)	0.75 (0.65, 0.87)	0.95 (0.91, 0.99)
South West	0.99 (0.98, 1.01)	0.61 (0.56, 0.66)	1.22 (1.01, 1.47)	0.60 (0.56, 0.65)
York	0.89 (0.88, 0.90)	0.88 (0.83, 0.93)	1.03 (0.88, 1.20)	0.77 (0.72, 0.81)
Underlying chronic health conditions				
Asthma	1.09 (1.09, 1.10)	0.86 (0.83, 0.89)	0.85 (0.78, 0.93)	0.93 (0.89, 0.96)
COPD	1.23 (1.21, 1.24)	0.89 (0.82, 0.96)	0.83 (0.70, 0.98)	1.05 (0.97, 1.13)
Hypertension	0.98 (0.98, 0.99)	1.12 (1.08, 1.16)	1.10 (1.00, 1.20)	1.13 (1.09, 1.18)
Diabetes	0.98 (0.97, 0.99)	1.26 (1.21, 1.31)	1.18 (1.07, 1.30)	1.19 (1.15, 1.24)
Congestive heart failure	1.26 (1.24, 1.28)	0.99 (0.92, 1.07)	0.92 (0.77, 1.11)	1.25 (1.16, 1.35)

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

Dementia or frailty score >15	2.12 (2.09, 2.15)	1.38 (1.29, 1.49)	1.69 (1.44, 1.98)	2.59 (2.41, 2.79)
Cancer[c]	1.13 (1.12, 1.15)	0.72 (0.66, 0.80)	0.87 (0.70, 1.09)	0.85 (0.78, 0.94)
Chronic kidney disease[c]	1.31 (1.29, 1.32)	0.86 (0.80, 0.93)	1.04 (0.88, 1.23)	1.16 (1.08, 1.24)
Immunocompromised[d]	1.30 (1.28, 1.33)	0.79 (0.70, 0.89)	0.84 (0.63, 1.10)	1.02 (0.91, 1.14)
Advanced liver disease	1.14 (1.11, 1.16)	0.78 (0.68, 0.90)	0.76 (0.54, 1.07)	0.87 (0.76, 1.00)
Ischemic heart disease	1.02 (1.01, 1.03)	0.88 (0.82, 0.95)	0.81 (0.68, 0.97)	0.90 (0.84, 0.97)
Ischemic stroke or TIA[e]	1.15 (1.13, 1.17)	1.05 (0.95, 1.17)	0.94 (0.73, 1.20)	1.18 (1.06, 1.31)
Schizophrenia[f]	1.24 (1.21, 1.27)	0.84 (0.73, 0.96)	1.28 (0.92, 1.77)	0.99 (0.86, 1.13)
Substance abuse[f]	1.17 (1.16, 1.19)	0.72 (0.65, 0.79)	0.52 (0.40, 0.68)	0.86 (0.78, 0.94)
Healthcare use				
Aggregated Diagnostic Group quintile (ref: 0 ADGs)				
2 (1-2 ADGs)	1.64 (1.62, 1.66)	0.78 (0.74, 0.84)	0.68 (0.58, 0.81)	1.24 (1.17, 1.32)
3 (3-4 ADGs)	2.07 (2.04, 2.10)	0.75 (0.70, 0.80)	0.57 (0.48, 0.68)	1.47 (1.37, 1.57)
4 (5-6 ADGs)	2.45 (2.41, 2.48)	0.74 (0.69, 0.79)	0.56 (0.47, 0.68)	1.69 (1.58, 1.82)
5 (7-27 ADGs)	3.11 (3.06, 3.15)	0.75 (0.69, 0.80)	0.52 (0.43, 0.63)	2.13 (1.98, 2.29)
Hospital admissions, past 3 years (ref: 0 admissions)				
1	0.99 (0.98, 1.00)	0.93 (0.89, 0.97)	0.88 (0.79, 0.98)	0.92 (0.88, 0.96)
2	1.16 (1.14, 1.17)	0.86 (0.79, 0.93)	0.80 (0.66, 0.97)	1.01 (0.94, 1.09)
≥3	1.75 (1.73, 1.78)	0.82 (0.76, 0.90)	1.04 (0.86, 1.26)	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)	1.07 (0.95, 1.19)	1.12 (1.08, 1.18)
5 – 8	1.11 (1.10, 1.12)	1.03 (0.98, 1.08)	1.01 (0.89, 1.15)	1.17 (1.11, 1.23)
9 – 14	1.17 (1.16, 1.18)	0.98 (0.93, 1.04)	1.04 (0.90, 1.20)	1.17 (1.10, 1.24)
≥15	1.36 (1.34, 1.37)	0.91 (0.85, 0.97)	0.88 (0.75, 1.03)	1.22 (1.14, 1.30)
Influenza vaccination (2019-2020 season)	1.08 (1.08, 1.09)	0.81 (0.78, 0.83)	0.86 (0.79, 0.93)	0.87 (0.84, 0.90)
Environmental determinants[g]				
$PM_{2.5}$ category (μg/m^3) (ref: 2 – 6)				
6-7	0.97 (0.96, 0.98)	0.91 (0.84, 0.98)	0.81 (0.69, 0.96)	0.92 (0.85, 0.99)
7-8	0.91 (0.90, 0.92)	1.05 (0.95, 1.16)	0.99 (0.80, 1.22)	0.96 (0.87, 1.06)
8-9	0.90 (0.89, 0.92)	1.24 (1.11, 1.37)	1.15 (0.93, 1.43)	1.15 (1.04, 1.27)
≥10	0.89 (0.88, 0.91)	1.37 (1.22, 1.55)	1.49 (1.18, 1.90)	1.24 (1.10, 1.39)
NO_2 category (parts per billion) (ref: 0 – 6)				
6-8	0.96 (0.95, 0.96)	1.05 (1.00, 1.11)	1.05 (0.94, 1.16)	1.00 (0.95, 1.05)
≥8	0.93 (0.92, 0.95)	1.12 (1.05, 1.19)	1.08 (0.93, 1.26)	1.03 (0.97, 1.10)
Social determinants of health[h] (area-level)				
Household income quintile (ref: 1st quintile,				

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

lowest income)				
2	1.03 (1.02, 1.04)	0.98 (0.93, 1.03)	0.95 (0.84, 1.08)	1.02 (0.97, 1.06)
3	1.09 (1.08, 1.10)	1.06 (0.99, 1.12)	0.96 (0.82, 1.12)	1.13 (1.07, 1.20)
4	1.11 (1.10, 1.13)	1.02 (0.95, 1.09)	0.95 (0.80, 1.12)	1.09 (1.02, 1.17)
5 (highest income)	1.12 (1.10, 1.14)	1.00 (0.92, 1.08)	1.22 (1.01, 1.47)	1.08 (1.00, 1.16)
Social assistance quintile[i] (ref: 1st quintile)				
2	1.03 (1.02, 1.04)	1.11 (1.06, 1.16)	1.04 (0.93, 1.16)	1.12 (1.07, 1.18)
3	1.04 (1.04, 1.05)	1.19 (1.14, 1.26)	1.15 (1.02, 1.29)	1.22 (1.16, 1.28)
4	1.05 (1.04, 1.06)	1.20 (1.14, 1.27)	1.09 (0.96, 1.25)	1.22 (1.16, 1.30)
5	1.05 (1.04, 1.06)	1.36 (1.27, 1.45)	1.11 (0.95, 1.30)	1.38 (1.30, 1.47)
Limited educational attainment quintile[j] (ref: 1st quintile)				
2	1.04 (1.03, 1.04)	1.11 (1.06, 1.16)	1.13 (1.01, 1.26)	1.15 (1.10, 1.20)
3	1.02 (1.02, 1.03)	1.23 (1.17, 1.29)	1.21 (1.08, 1.36)	1.26 (1.21, 1.33)
4	1.03 (1.02, 1.04)	1.33 (1.27, 1.40)	1.46 (1.29, 1.65)	1.36 (1.30, 1.43)
5	1.02 (1.01, 1.03)	1.49 (1.41, 1.58)	1.50 (1.31, 1.72)	1.52 (1.44, 1.60)
Uncoupled quintile[k] (ref: 1st quintile)				
2	1.02 (1.01, 1.03)	0.96 (0.91, 1.01)	1.06 (0.96, 1.18)	0.97 (0.93, 1.02)
3	1.07 (1.06, 1.08)	0.96 (0.92, 1.01)	1.00 (0.88, 1.12)	1.01 (0.96, 1.06)
4	1.20 (1.18, 1.21)	0.98 (0.93, 1.04)	0.95 (0.83, 1.09)	1.14 (1.08, 1.21)
5	1.39 (1.38, 1.41)	1.09 (1.03, 1.17)	1.06 (0.91, 1.25)	1.46 (1.37, 1.56)
Household density quintile[l] (ref: 1st quintile)				
2	1.02 (1.01, 1.02)	1.25 (1.19, 1.32)	1.24 (1.10, 1.40)	1.25 (1.19, 1.32)
3	1.04 (1.03, 1.05)	1.48 (1.40, 1.57)	1.36 (1.18, 1.57)	1.53 (1.45, 1.62)
4	1.03 (1.02, 1.04)	1.84 (1.74, 1.94)	1.82 (1.59, 2.09)	1.89 (1.79, 1.99)
5	0.98 (0.97, 0.99)	2.12 (1.99, 2.26)	1.74 (1.47, 2.05)	2.08 (1.95, 2.21)
Apartment building density category[m] (ref: 1st category)				
2	1.03 (1.03, 1.04)	0.98 (0.94, 1.03)	1.00 (0.90, 1.10)	1.00 (0.96, 1.04)
3	1.02 (1.01, 1.03)	1.13 (1.08, 1.20)	1.08 (0.94, 1.24)	1.16 (1.10, 1.22)
Recent immigration category[n] (ref: 1st category)				
2	1.00 (0.99, 1.00)	1.05 (1.00, 1.09)	1.12 (1.02, 1.23)	1.04 (1.00, 1.08)
3	0.94 (0.94, 0.95)	1.18 (1.12, 1.23)	1.31 (1.17, 1.47)	1.12 (1.07, 1.16)
Visible minority quintile[o] (ref: 1st category)				
2	0.99 (0.98, 0.99)	0.94 (0.88, 1.01)	1.02 (0.90, 1.16)	0.94 (0.88, 1.00)
3	0.94 (0.93, 0.95)	0.94 (0.88, 1.01)	1.04 (0.91, 1.20)	0.89 (0.83, 0.96)
4	0.91 (0.90, 0.92)	1.03 (0.96, 1.11)	1.10 (0.94, 1.29)	0.94 (0.87, 1.01)
5	0.85 (0.84, 0.87)	1.22 (1.12, 1.33)	1.15 (0.95, 1.39)	1.03 (0.95, 1.12)

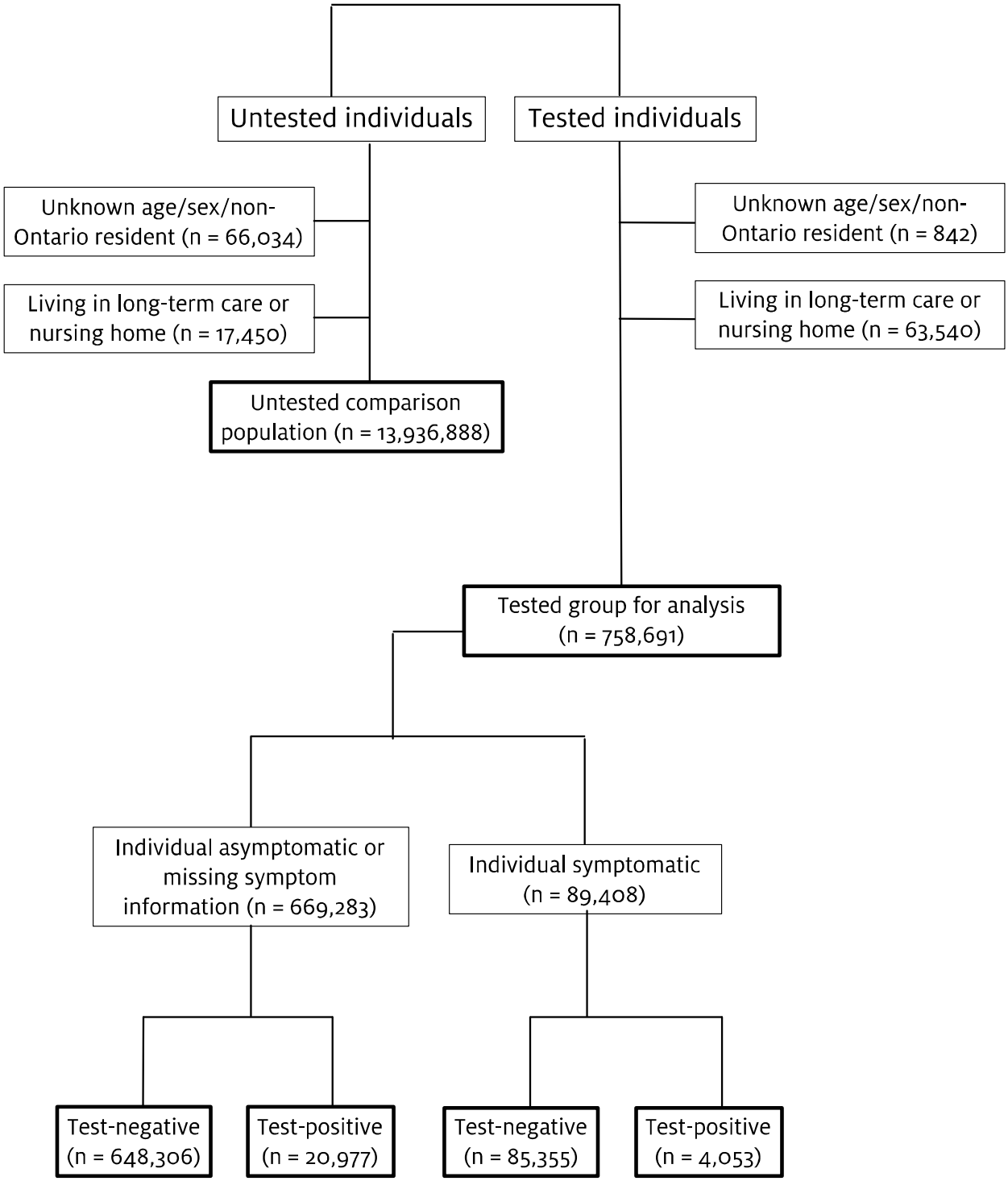
medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[a] Fully-adjusted analyses contain as covariates all variables listed in this table.
"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.[80]
[g] All variables in this category are area-level variables at the level of the Census Dissemination Area.
[h] 1st quintile represents 38.0 – 63.5% of individuals in the area receiving governmental transfer payments; 2nd quintile: 63.5 – 67.2% of individuals; 3rd quintile: 67.2 – 70.8% of individuals; 4th quintile: 70.8 - 75.3% of individuals; 5th quintile: 75.3 – 98.1% of individuals.
[i] 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.
[j] 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.
[k] 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.
[l] 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.
[m] 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.
[n] 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.

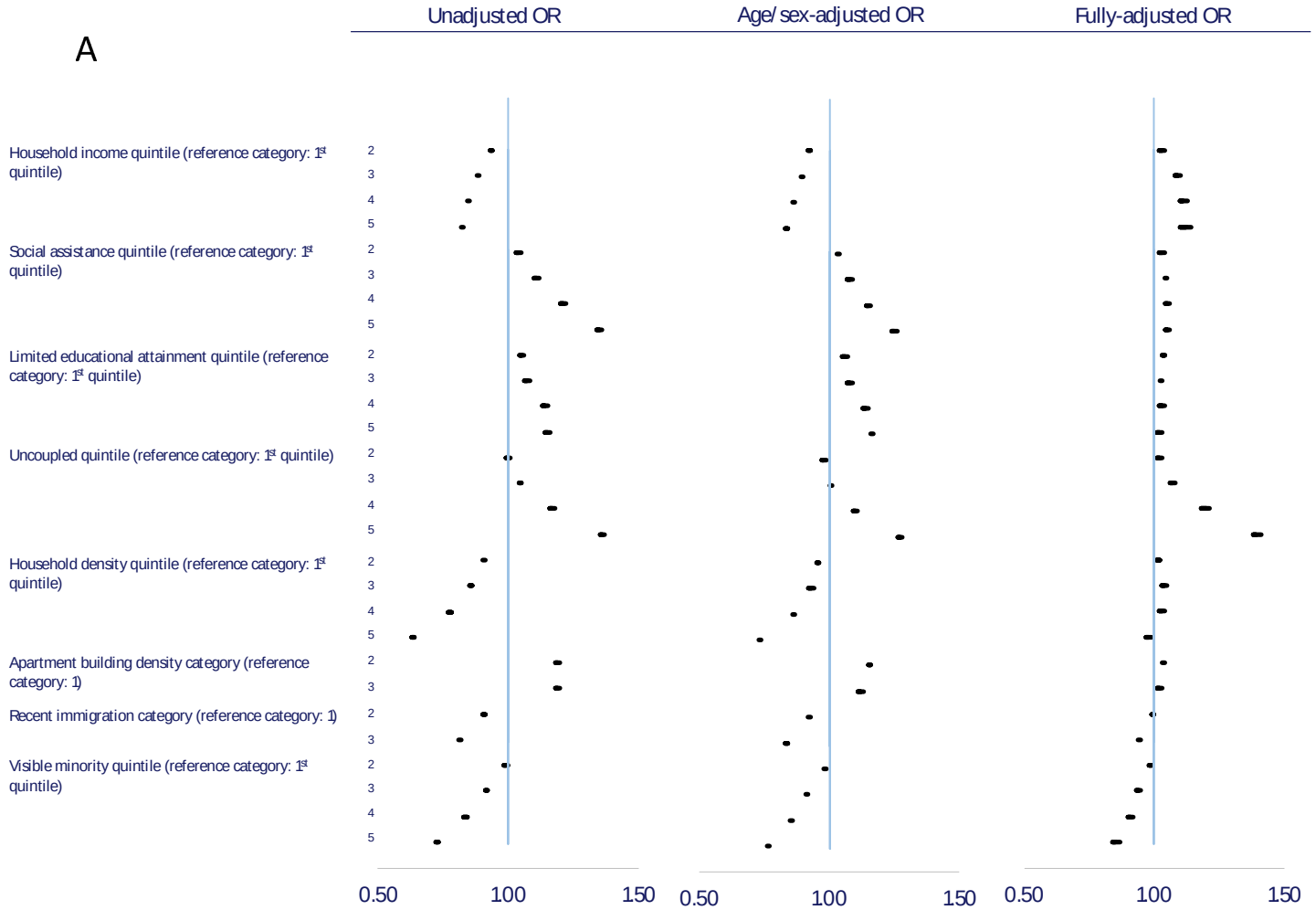
medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

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



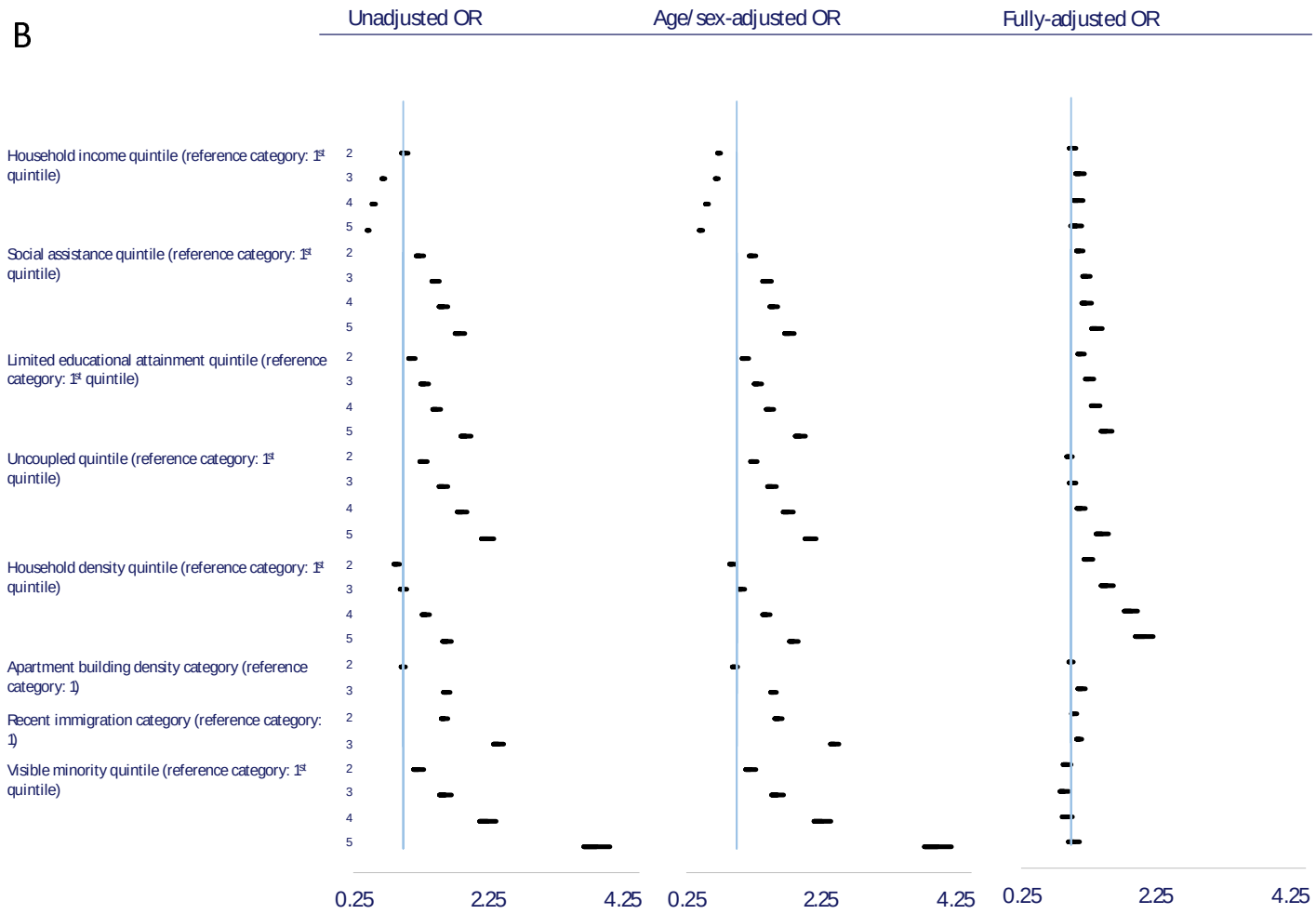
Individuals were excluded if they had a listed age >105 years, had a listed postal code outside of Ontario, or had no record of contact with the healthcare system in the past 8 years (since March 2012).

Figure 2. Unadjusted, sex/age adjusted, and fully adjusted association between social determinants and SARS-CoV-2 testing (panel A) and 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. Age/sex-adjusted models are shown in Table 2. Fully-adjusted models control for all variables listed in Table 2, including (but not limited to) all variables shown in this figure; effect estimates and 95% CIs for all models shown here can be found in eTable 2.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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 (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. Effect estimates and 95% CIs for all models shown here can be found in eTable 5.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.09.20223792; this version posted November 12, 2020. 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.