Coronavirus Disease 2019 (COVID-19) testing, infection and complication rates among individuals at risk of homelessness in Ontario, Canada: A retrospective cohort study

Journal:	*CMAJ*
Manuscript ID	CMAJ-20-1773
Manuscript Type:	Research - Cohort study (retrospective)
Date Submitted by the Author:	23-Jul-2020
Complete List of Authors:	Richard, Lucie; Institute for Clinical Evaluative Sciences, ICES Western Booth, Richard; Western University, Arthur Labatt Family School of Nursing Clemens, Kristin; Western University, Medicine Forchuk, Cheryl; Western University Shariff, Salimah; Institute for Clinical Evaluative Sciences, ICES Western
Keywords:	Epidemiology, Infectious Diseases, Public Health
More Detailed Keywords:	Homelessness
Abstract:	Background: Individuals at risk of homelessness (IARH) are believed to be at high risk of COVID-19 infection and, when infected, of experiencing complications. We aimed to describe and compare COVID-19 testing, positivity, hospitalization and mortality rates as of May 31, 2020 for IARH vs. community-dwelling individuals. Methods: We conducted a population-based retrospective cohort study in Ontario, Canada using linked health administrative data among individuals eligible for healthcare who are either a) IARH (n=27,671) or b) community-dwelling non-IARH (n=14,485,052). We examined COVID-19 testing, positive COVID-19 tests, hospitalization for COVID-19 and death within 21 days of a positive test from multivariable cox proportional hazard models. Results: As of May 31 2020, 11.68% IARH received at least one test for COVID-19, compared to 2.33% community-dwelling individuals (adjusted hazard ratio [aHR] 3.88, 95% CI 3.74-4.02). COVID-19 positivity was 5.57% among IARH, compared to 5.84% among community-dwelling individuals (aHR 1.14, 95% CI 0.98-1.32). IARH were more likely to be hospitalized for confirmed COVID-19 (aHR 14.41, 95% CI 10.99-18.90) and to experience death within 21 days of a positive test (aHR 10.79, 95% CI 4.35-26.76). Interpretation: In Ontario, IARH were more likely to be tested for COVID-19 than the community-dwelling population. Though the overall positivity was higher, case rates were similar once accounting for testing rates, possibly thanks to screening and sheltering countermeasures

	implemented in many cities. Despite this, IARH had much higher COVID-19 related hospitalization and death rates when infected, even after controlling for other known risk factors.

Coronavirus Disease 2019 (COVID-19) testing, infection and complication rates among individuals at risk of homelessness in Ontario, Canada

Richard, Lucie, MA, ICES Western
Booth, Richard, PhD, Western University
Jennifer Rayner, PhD, Alliance for Healthier Communities
Clemens, Kristin K, MD, MSc, Western University
Forchuk, Cheryl, PhD, Western University
Shariff, Salimah Z., PhD, ICES Western

Corresponding author: Lucie Richard lucie.richard@ices.on.ca

Funding Statement: This study was supported by ICES Western.

Declaration of author(s) competing interests: Outside of the submitted work, Dr. Clemens received a diabetes research award funded in part by Astra Zeneca. Dr. Clemens has attended conferences sponsored by Merck Inc. She has received honoraria for delivering Certified Medical Education talks from Sutherland Global Services and the Toronto Ontario Knowledge Translation Working Group Inc.

Author contributions: LR conceived of the study, participated in the study coordination, study design, acquisition of data and interpretation of results, performed the analysis and drafted the manuscript. SZS conceived of the study, participated in the study design, interpretation of study results and provided feedback on the manuscript. RB participated in the study design, acquisition of data, and interpretation of study results and provided feedback on the manuscript. CF, JR and KC contributed to the study design, interpretation of study results and provided feedback on the manuscript. All authors read and approved the final manuscript.

Word count: 2,498

Keywords: Homelessness; health services research; COVID-19

Abstract

Background: Individuals at risk of homelessness (IARH) are believed to be at high risk of COVID-19 infection and, when infected, of experiencing complications. We aimed to describe and compare COVID-19 testing, positivity, hospitalization and mortality rates as of May 31, 2020 for IARH vs. community-dwelling individuals.

Methods: We conducted a population-based retrospective cohort study in Ontario, Canada using linked health administrative data among individuals eligible for healthcare who are either a) IARH (n=27,671) or b) community-dwelling non-IARH (n=14,485,052). We examined COVID-19 testing, positive COVID-19 tests, hospitalization for COVID-19 and death within 21 days of a positive test from multivariable cox proportional hazard models.

Results: As of May 31 2020, 11.68% IARH received at least one test for COVID-19, compared to 2.33% community-dwelling individuals (adjusted hazard ratio [aHR] 3.88, 95% CI 3.74-4.02). COVID-19 positivity was 5.57% among IARH, compared to 5.84% among community-dwelling individuals (aHR 1.14, 95% CI 0.98-1.32). IARH were more likely to be hospitalized for confirmed COVID-19 (aHR 14.41, 95% CI 10.99-18.90) and to experience death within 21 days of a positive test (aHR 10.79, 95% CI 4.35-26.76).

Interpretation: In Ontario, IARH were more likely to be tested for COVID-19 than the community-dwelling population. Though the overall positivity was higher, case rates were similar once accounting for testing rates, possibly thanks to screening and sheltering countermeasures implemented in many cities. Despite this, IARH had much higher COVID-19 related hospitalization and death rates when infected, even after controlling for other known risk factors.

Introduction

In the early months of the COVID-19 pandemic, significant discourse emerged about the theoretical or estimated impact of the virus on homeless individuals (1–5). Individuals experiencing homelessness are not only believed to be at increased risk of COVID-19 infection, due to high population density in emergency shelters and other precarious housing conditions (3) but are also at higher risk of negative outcomes following infection given the high level of comorbidities experienced by this population (6). Indeed, by May 2020 numerous shelters in Canadian and American cities reported COVID-19 outbreaks (7-9), prompting sheltering and distancing countermeasures in many Ontario cities (e.g. moving people into hotels) (10-18). Capturing the true number of individuals experiencing homelessness at a given time is particularly difficult (19) and as such there remains little understanding of the epidemiology of COVID-19 among this population in Canada. There is also no evidence to date assessing how their risk of COVID-19 and its complications differs from that of the general, community dwelling population.

Using a recently validated case definition for risk of homelessness (20) and population-level health administrative databases, our aim was to describe and compare COVID-19 testing, positivity, hospitalization and death rates for individuals at risk of homelessness (IARH) compared to community-dwelling individuals not identified as IARH as of May 31 2020 in Ontario, Canada.

Methods

Study design and setting

We conducted a retrospective population-based cohort study in Ontario, Canada's most populous province (21) using health administrative data. Databases were linked using unique encoded identifiers and analyzed at ICES (formerly known as the Institute for Clinical Evaluative Sciences) (22), a not-for-profit research institute. ICES is a prescribed entity under section 45 of Ontario's Personal Health Information Protection Act, which authorizes ICES to collect personal health information, without consent, for the purpose of analysis or compiling statistical information with respect to the management of, evaluation or monitoring of, the allocation of resources to or planning for all or part of the health system. 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. This study follows RECORD reporting guidelines (23).

Data sources

Health administrative databases used to define participants, outcome and covariates included the a) Canadian Institute for Health Information (CIHI) Discharge Abstract Database (DAD) and Same Day Surgery (SDS) databases; b) National Ambulatory Care Reporting System emergency (NACRS); c) Ontario Mental Health Reporting System (OMHRS); d) Ontario Laboratories Information System (OLIS), e) ICES Registered Persons Database Demographic and Postal Year Datasets, f) Ontario Health Insurance Physicians (OHIP) claims database, g) Immigration, Refugee and Citizenship Canada (IRCC) Permanent Residents database, h) Ontario Drug Benefit (ODB) database, and several ICES-derived population-surveillance datasets including: the Chronic Obstructive Pulmonary Disease (COPD)(24), Ontario Asthma Dataset (25), Ontario Diabetes Dataset (ODD)(26), Congestive Heart Failure (CHF)(27), Ontario Hypertension Dataset (HYPER)(28) and Ontario Rheumatoid Arthritis Dataset (ORAD)(29) derived cohorts. Details about the databases used in participant selection, description and outcome ascertainment are presented in Supplement Tables 1 and 2.

Participants

Participants were comprised of two groups, followed from January 23 2020, the date of the first known COVID-19 case in Ontario (30), until May 31 2020 for evidence of COVID-19 lab testing. The community-dwelling comparison group included Ontario residents eligible for health coverage who were not living in an institutionalized facility (long-term care, hospitalized or receiving continuing care services in hospital) and not identified as IARH (defined below) as of January 23 2020. We identified individuals at risk of homelessness (IARH), as anyone eligible for health coverage in Ontario who were not living in an institutionalized facility as of January 23 2020, and who met the case definition of homelessness during a healthcare encounter between October 1 2018 and May 31 2020 (case definition summarized in Supplement Table 1). The selected case definition was adopted from a recent validation study (20); we *a priori* elected to extend the case definition to May 31 2020 to more comprehensively capture individuals identified as homeless during the

pandemic. In the previous validation work, sensitivity of the selected algorithm was found to be 33.2% with specificity over 99.9% (20). However, this validation was conducted prior to April 2018, when Canadian Institute for Health Information (CIHI) mandated the reporting of homelessness using the International Classification of Diseases, 10th revision (ICD-10) codes Z59.0 and Z59.1 within Canadian hospitals (31). We anticipate this change in coding practice will have increased the sensitivity of the case definition.

Outcome Measure

Our primary outcome measures were the receipt and result of a COVID-19 test, as recorded in the Ontario Laboratory Information System (OLIS). COVID-19 test results were categorized as positive or non-positive (ie. negative), with indeterminate and pending tests considered non-positive. Cancelled or rejected tests were categorized as non-tested. Where multiple tests records were present, an individual was deemed positive if any test indicated a positive result. Primary outcomes were measured for each group provincially as well as by Local Health Integration Network (LHIN). In Ontario, until recently, coordination of care was organized within geographically defined LHINs (32). Our secondary outcomes of interest included a) hospitalization with ICD-10 code U07.1 (coronavirus, positive) following a positive test and b) death within 21 days of the first positive COVID-19 test.

Other Covariates

We obtained characteristics of participants at baseline (January 23 2020) including age, sex, neighbourhood income quintile, Ontario Marginalization Index (ON-Marg) subscales (deprivation, dependency, instability, and ethnic concentration)(33), level of urbanicity, immigrant and refugee status, region of residence (LHIN), Charlson comorbidity index, presence of comorbidities (including asthma, chronic obstructive pulmonary disorder (COPD), diabetes, congestive heart failure, hypertension, chronic liver disease and rheumatoid arthritis), number of recent primary healthcare encounters (as an indicator for healthcare usage), recent healthcare encounters for mental health related issues and enrolment with a family doctor (variable definitions in Supplement Table 3).

Analysis

We conducted all analyses using SAS, version 9.4 (34). We compared baseline covariates between groups using one-way ANOVA, Kruskal-Wallis, Cochran-Armitage or Chi-square tests, as appropriate. Due to the large sample size we also report standardized differences, which assess differences between group means as a percentage of the pooled standard deviation; a difference of less than 10% was considered insignificant (35). We calculated crude outcome rates for COVID-19 testing and positivity at the provincial and regional levels. We used multivariable Cox Proportional Hazards models to quantify the risk (hazard) of experiencing the outcome for the IARH group, using the community-dwelling population as a comparator. Censoring events included entry into long term care (no longer community-dwelling), death (not applicable for the mortality outcome), or the maximum available follow-up for which data were available at the time of the analysis (May 31 2020). Model results are presented as unadjusted and adjusted hazard ratios (aHR), with accompanying 95% confidence interval (CI).

Results

We identified 27,671 IARH and 14,485,052 community-dwelling non-IARH (Figure 1). IARH were younger (less than 40 years of age: 52.1% vs 47.3%), male (67.9% vs. 49.2%), less likely to have immigrated to Canada (1.8% vs. 4.3%), reside in the lowest income (43.4% vs. 18.8%) and most marginalized neighbourhoods (highest deprivation quintile: 38.8% vs 17.8%; highest instability quintile: 50.7% vs 22.1%), respectively (Table 1). Most IARH lived in large census metropolitan areas (CMAs)(58.2%), but compared to non-IARH, more lived in small to medium census CMAs (34.1% vs 27.3%) and non-CMA areas. IARH had higher comorbidity (charlson 2+: 5.8% vs 1.5%) and higher rates of respiratory conditions (Asthma: 15.6% vs 11.2%; COPD:12.1% vs 6.2%) and chronic liver disease (14.7% vs 2.3%). They were also greater users of primary healthcare (median: 20 visits [IQR 7-49] vs 5 [IQR 1-11]), and of mental healthcare (psychosis: 33.0% vs 1.2%; non-psychosis: 57.4% vs 12.0%; substance-abuse related: 52.4% vs 1.4%). Finally, enrolment with a primary care physician was low (48.4% vs 75.2%).

From January 23 to May 31, 3,232 or 11.68% of IARH received at least one COVID-19 test (Table 2), compared to 337,448 (or 2.33%) non-AHRI (unadjusted HR 5.32 [95% CI 5.14-5.51]). Regionally, testing rates by LHIN ranged between a low of 7.99% to a high of 13.84% for IARH (figure 2a), and between a low of 1.87% and a high of 3.04% for non-IARH. After adjusting for age, sex, income, ON-Marg deprivation and instability subscales, immigrant status, level of urbanicity, presence of comorbidities, outpatient healthcare usage and enrollment with a family doctor, the hazard of being tested was 3.88 (CI: 3.74-4.02).

We identified 180 IARH (5.57% of tested persons, or 0.65% overall) with a positive COVID-19 test result compared to 19,701 (5.84% of tested persons or 0.14% overall) for non-IARH (unadjusted HR 1.09 [95% CI 0.94-1.26])(table 2). At the regional level, positivity among tested ranged from a low of 0% (no cases) to a high of 11.40% among IARH (Figure 2b), and 0.98% to a high of 12.89% among non-IARH. Positivity in most regions was similar or lower among IARH (values under 1)(Figure 2c). The hazard ratio of receiving a positive test among those tested remained non-significant after adjustment (aHR 1.14 [95% CI 0.98-1.32]).

Finally, we identified 54 IARH (30% of cases, 1.67% of tested individuals and 0.2% overall) who were hospitalized with confirmed COVID-19 compared to 352 non-IARH (8.64% of cases, 0.5% of tested individuals and 0.01% overall)(table 2). After adjustment, IARH had a 14-fold rate of hospitalization (aHR 14.41 [95% CI 10.99-18.90]). While we are unable to report the absolute number of deaths to protect patient privacy, IARH had an over 10-fold rate of death within 21 days of a positive COVID-19 test, even after adjustment (aHR 10.79 [95% CI 4.35-26.76]).

Discussion

This study offers the first population-level assessment of the impact of COVID-19 on IARH. We found that testing rates were much higher in IARH compared to community-dwelling Ontarians. For much of the study period, Ontario primarily tested for COVID-19 in individuals presenting with symptoms who had recent out-of-country travel or pre-existing risk factors (ie. advanced age; significant comorbidities). These restrictions did not apply to individuals experiencing homelessness, who received testing priority status (36-38). Thus, advocacy and policy at least partially contributed to this finding. In turn, the high testing rate likely contributed to higher positivity among IARH (0.65% vs 0.14%). Within the

subset of individuals who were tested, however, IARH were not more likely to test positive, even after controlling for important risk factors. Similarly, the ratio of positivity was equal or favourable to IARH in most health regions. These findings may serve as early, untested evidence that sheltering countermeasures implemented in many Ontario cities to prevent outbreaks (12-18) may have been effective in protecting IARH against COVID-19 infection. Finally, IARH were more likely to be hospitalized for COVID-19 and experience death within 21 days of a positive test, even after accounting for underlying risk factors. Part of the heightened rate of hospitalization may reflect policies in certain areas placing IARH cases in hospitals for isolation when no alternative shelter is available (39-40); and part of the heightened death rate reflects the heightened mortality experienced by the homeless generally (41-42). However, despite these exacerbating factors, our results indicate a clear heightened risk for COVID-19 related complications among IARH.

Few studies have detailed the epidemiological impact of COVID-19 on populations experiencing homelessness. Two reports detail within-homeless case rates in shelters of several American cities, with positivity among tested ranging between 5.7% and 18% (9, 43), but do not compare to the general population. Another report in Boston followed COVID-19 cases over a 15-day period, identifying a cumulative case rate of 4.6% among the homeless vs. 0.19% in the general population (8). The Boston report's findings are not directly comparable to our findings for several reasons, though. First, the report used, likely due to lack of alternatives, point-in-time (PIT) counts to estimate the homeless population denominator. However, PIT counts are not meant for use as an absolute population estimate and have been found to be unreliable when used for this purpose (44). Second, positivity rates were generated using the entire homeless population as the denominator, rather than within-tested individuals, as would be more appropriate when bias in testing between groups exists. Finally, the health care system in the United States differs substantially from Ontario's government-funded health care system where a requirement to pay for care may influence testing rates and, in turn, positivity rates.

Our study is not without limitations. First, we could only include individuals eligible for Ontario health care coverage, which although near-universal (>99%) does not include recent arrivals to Ontario, Indigenous persons on reserves, certain non-convention refugee claimants, certain veterans and serving members of the Canadian Forces (45). As Indigenous persons and refugees are over-represented in Canada's homeless population, our counts are likely underestimates, particularly in the Greater Toronto Region, where refugees consist of one third of shelter users (46).

Second, as homelessness is not comprehensively tracked in our province, we applied an algorithm, developed by authors of this report, that leverages hospital-based administrative data (20). Our previous validation demonstrated that these algorithms, although highly specific have moderate to poor sensitivity, largely due to the transiency of homelessness. Thus, despite recent changes to CIHI policies making reporting of homelessness mandatory (31), our cohort is likely smaller than the true Ontario homeless population. However, characteristics of our identified population (Table 1) are similar to other reports of homeless populations in Canada (19, 20). Therefore, we do not believe our undercount biases our rates. However, as previously mentioned, results can only be generalized to individuals with health care coverage in Ontario. Finally, despite our attempts to control for risk factors, comparator groups had different conditions for COVID-19 testing, and therefore complication rates should be interpreted with caution.

Legitimate concerns exist about the potential impact of the COVID-19 pandemic on individuals experiencing homelessness. In Ontario, we found that priority testing status and screening efforts by shelters and municipalities may have resulted in significantly higher COVID-19 testing (and thus positive cases) rates among IARH. We also found that IARH infected with COVID-19 experienced significantly higher rates of hospitalization for COVID-19 and death following a positive COVID-19 test. Future work should update this short-term evaluation, particularly once homeless sheltering countermeasures are relaxed, and more directly assess the impact of sheltering countermeasures on infection and complication rates.

Acknowledgements

This study was conducted and supported by the ICES Western Site, which receives core funding from the Academic Medical Organization of Southwestern Ontario (AMOSO), the Schulich School of Medicine and Dentistry, Western University, and the Lawson Health Research Institute. ICES is funded by an annual grant from the Ontario Ministry of Health. The opinions, results, and conclusions reported are those of the authors and are independent from the funding sources. No endorsement by ICES or the funding agencies is intended or should be inferred. Parts of this material are based on data and/or information compiled and provided by CIHI. However, the analyses, conclusions, opinions and statements expressed in the material are those of the author(s), and not necessarily those of CIHI. Information in the

Immigration, Refugees and Citizenship Canada Permanent Resident Dataset were provided by Immigration, Refugees and Citizenship Canada.

References

1. Patton J. (2020, April 21) Coronavirus: Advocates threaten legal action if Toronto doesn't step up aide for homeless population. *Global News*. Retrieved from https://globalnews.ca/news/6847669/coronavirus-toronto-homeless-aid/
2. Volkow ND (2020). Collision of the COVID-19 and Addiction Epidemics. *Ann Intern Med*. 2019:2019-2021. doi:10.7326/m20-1212
3. Tsai J, Wilson M. Comment COVID-19 : a potential public health problem for homeless populations. *Lancet Public Heal*. 2020;2(20):19-20. doi:10.1016/S2468-2667(20)30053-0
4. Culhane D, Treglia D, Steif K, Kuhn R, & Byrne T. (2020). Estimated emergency and observational/quarantine capacity need for the US homeless population related to COVID-19 exposure by county; projected hospitalizations, intensive care units and mortality.
5. Lima NNR, de Souza RI, Feitosa PWG, de Sousa Moreira JL, da Silva CGL, & Neto MLR (2020). People experiencing homelessness: Their potential exposure to COVID-19. *Psychiatry Research*, 112945.
6. Fazel S, Geddes JR, Kushel M. The health of homeless people in high-income countries: descriptive epidemiology, health consequences, and clinical and policy recommendations. *Lancet*. 2014;384(9953):1529-1540. doi:10.1016/S0140-6736(14)61132-6
7. Casey L (2020, April 14). COVID-19 spreading among Toronto's homeless, with 30 cases reported. *Canadian Broadcasting Corporation.* Retrieved from https://www.cbc.ca/news/canada/toronto/homeless-population-toronto-covid-19-1.5532300
8. Baggett TP, Keyes H, Sporn N, & Gaeta JM (2020). Prevalence of SARS-CoV-2 infection in residents of a large homeless shelter in Boston. *JAMA*, 323(21), 2191-2192.
9. Mosites E, Parker EM, Clarke KE, Gaeta JM, Baggett TP, Imbert E, ... & Gonzales E. (2020). Assessment of SARS-CoV-2 infection prevalence in homeless shelters—four US cities, March 27–April 15, 2020. *Morbidity and Mortality Weekly Report*, 69(17), 521.
10. Yu A (2020, April 15). People are sleeping in the streets or ravines because shelters are unsafe": This doctor is helping restructure homeless shelters during the Covid crisis. *Toronto Life.* Retrived from https://torontolife.com/city/people-are-sleeping-in-the-streets-or-ravines-because-shelters-are-unsafe-this-doctor-is-helping-restructure-homeless-shelters-during-the-covid-crisis/.
11. Hanson H (2020, April 21). Ottawa Will Consider Buying Empty Hotels To House The City's Homeless People. *Narcity*. Retrieved from https://www.narcity.com/news/ca/on/ottawa/hotel-for-homeless-people-in-ottawa-is-being-considered-by-the-city.
12. D'Avino C (2020, April 5). Sault motel to offer emergency self-isolation rooms for city's less-fortunate. *CTV News Northern Ontario*. Retrieved from https://northernontario.ctvnews.ca/sault-motel-to-offer-emergency-self-isolation-rooms-for-city-s-less-fortunate-1.4883278
13. Bogdan S (2020, April 1). City of London moving homeless into motels to combat spread of COVID-19. *Global News*. Retrieved from https://globalnews.ca/news/6764303/coronavirus-london-ont-homeless-motels/
14. LaFreche G (2020, March 18). Homeless shelters adapt to pandemic, prepare to house infected clients. *St Catharines Standard*. Retrieved from https://www.niagarafallsreview.ca/news-story/9910529-homeless-shelters-adapt-to-pandemic-prepare-to-house-infected-clients/

15. Connor B (2020, April 6). Sudbury's homeless shelter temporarily relocating to hotel. *CTV News Northern Ontario*. Retrieved from https://northernontario.ctvnews.ca/sudbury-s-homeless-shelter-temporarily-relocating-to-hotel-1.4884916
16. Nielsen K (2020, March 23). Coronavirus: Radisson will be isolation shelter for Waterloo homeless who contract COVID-19. *Global News*. Retrieved from https://globalnews.ca/news/6719914/coronavirus-waterloo-radisson-homeless/
17. La Grassa J (2020, April 9). FirstOntario Centre will become a shelter for Hamilton's homeless during COVID-19. *Canadian Broadcasting Corporation*. Retrieved from https://www.cbc.ca/news/canada/hamilton/homeless-population-shelters-covid-19-hamilton-1.5527529
18. Johnstone H (2020, April 29). City opens arena to give homeless men some space. *Canadian Broadcasting Corporation*. Retrieved from https://www.cbc.ca/news/canada/ottawa/physical-distancing-centre-ottawa-homeless-1.5549111
19. Gaetz S, Dej E, Richter T, Redman M. The State of Homelessness in Canada 2016. Toronto: Canadian Observatory on Homelessness Press.
20. Richard L, Hwang SW, Forchuk C, et al. Validation study of health administrative data algorithms to identify individuals experiencing homelessness and estimate population prevalence of homelessness in Ontario, Canada (2019). *BMJ Open*. 9(10):e030221. doi:10.1136/bmjopen-2019-030221
21. Statistics Canada (2020, July 7). Population estimates, quarterly. Table 17-10-0009-01 (formerly CANSIM 051-0005). Retrieved from https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000901
22. ICES (2020). www.ices.on.ca.
23. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM; RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med*. 2015;12(10):e1001885.
24. Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T (2009). Identifying individuals with physician diagnosed COPD in health administrative databases. *COPD* ; 6(5):388-394
25. Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T (2009). Identifying patients with physician-diagnosed asthma in health administrative databases. *Can Respir J* ;16:183-8
26. Hux JE, Ivis F, Flintoft V, Bica A. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm (2002). *Diabetes Care*; 25(3):512-516.
27. Schultz SE, Rothwell DM, Chen Z, Tu K (2013). Identifying cases of congestive heart failure from administrative data: a validation study using primary care patient records. *Chronic Dis Inj Can* ; 33(3):160-166.
28. Tu K, Lipscombe LL, Chen ZL, for the CHEP Outcomes Research Taskforce (2008). Prevalence and incidence of hypertension from 1995 to 2005: a population-based study. *CMAJ*. May 20;178(11):1429-35
29. Widdifield J, Bombardier C, Bernatsky S, Paterson JM, Green D, Young J, Ivers N, Butt DA, Jaakkimainen RL, Thorne JC, Tu K (2014). An administrative data validation study of the accuracy of algorithms for identifying rheumatoid arthritis: the influence of the reference standard on algorithm performance. *BMC musculoskeletal disorders.* Jun 23;15(1):216.
30. Government of Ontario (2020, January 25). Ontario Confirms First Case of Wuhan Novel Coronavirus. Retrieved from https://news.ontario.ca/mohltc/en/2020/01/ontario-confirms-first-case-of-wuhan-novel-coronavirus.html.
31. Canadian Institute for Health Information DAD Abstracting Manual, 2018—2019. Ottawa, ON: CIHI; 2018.
32. Government of Ontario (2019, November 13). Ontario Taking Next Steps to Integrate Health Care System. Retrieved from https://news.ontario.ca/mohltc/en/2019/11/ontario-taking-next-steps-to-integrate-health-care-system.html

33. Public Health Ontario (2018, October 9). Ontario Marginalization Index (ON-Marg). Retrieved from https://www.publichealthontario.ca/en/data-and-analysis/health-equity/ontario-marginalization-index.
34. SAS Enterprise Guide. V9.4. Cary, NC: SAS Institute Inc.
35. Austin, PC (2009). Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. *Communications in statistics-simulation and computation*, 38(6), 1228-1234.
36. Ontario Ministry of Health (2020, April 1). COVID-19 Guidance: Homeless Shelters. Retrieved from https://myrnao.ca/sites/default/files/attached_files/Homeless%20Shelters%20COVID-19%20Guidance%20Document%20-%20March%2031_2020_final_for%20translation.pdf
37. City of Toronto (2020, June 22). COVID-19 Response for People Experiencing Homelessness. Retrieved from https://www.toronto.ca/news/city-of-toronto-covid-19-response-for-people-experiencing-homelessness/.
38. Ottawa Public Health (ND). COVID-19 Testing Criteria. Retrieved from https://www.ottawapublichealth.ca/en/shared-content/covid-19-testing-criteria.aspx.
39. Bingley M (2020, April 10). Coronavirus: Toronto's homeless COVID-19 patients sent to hospital due to lack of recovery site. *Global News Toronto*. Retrieved from https://globalnews.ca/news/6806485/coronavirus-toronto-homeless-hospitals-covid-19-recovery-site/
40. Snider C (2020, June 17). How COVID-19 cases among the homeless overwhelm St. Mike's hospital. *Toronto Star*. Retrieved from https://www.thestar.com/opinion/contributors/2020/06/17/how-covid-19-cases-among-the-homeless-overwhelm-st-mikes-hospital.html
41. Morrison, DS (2009). Homelessness as an independent risk factor for mortality: results from a retrospective cohort study. *International Journal of Epidemiology*, Vol38, Issue 3, June 2009, Pg 877–883, https://doi.org/10.1093/ije/dyp160
42. Hwang SW, Wilkins R, Tjepkema M, O'Campo PJ, Dunn JR (2009). Mortality among residents of shelters, rooming houses, and hotels in Canada: 11 year follow-up study. *BMJ*; 339 :b4036
43. Tobolowsky FA, Gonzales E, Self JL, Rao CY, Keating R, Marx GE, McMichael TM, Lukoff MD, Duchin JS, Huster K, Rauch J, McLendon H … Kay, M (2020). COVID-19 Outbreak Among Three Affiliated Homeless Service Sites - King County, Washington, 2020. *Morbidity and mortality weekly report* 69(17), 523–526. https://doi.org/10.15585/mmwr.mm6917e2
44. National Alliance to End Homelessness (2014, February). 5 Myths About PIT Counts. Retrieved from https://www.homelessnesslearninghub.ca/sites/default/files/resources/5%20myths%20about%20Pit%20counts.pdf
45. Government of Ontario (2020, July 8). Apply for OHIP and get a health card. Retrieved from https://www.ontario.ca/page/apply-ohip-and-get-health-card
46. City of Toronto (2019, November 6). Refugees & Asylum Claimants Experiencing Homelessness in the City of Toronto. Retrieved from https://www.toronto.ca/news/refugees-asylum-claimants-experiencing-homelessness-in-the-city-of-toronto/

Table 1 – Cohort characteristics at index, by group

Characteristic at index	IARH[1] (N=27,671)	Community-dwelling Ontario population (N=14,485,052)	p-value	STD difference
Age, median (IQR)	38 (29-52)	41 (23-59)		5%
Age group, n (%)				
Youth (< 25 yrs)	4,474 (16.2%)	3,994,343 (27.6%)	<.001	28%
25 to 39 yrs	9,946 (35.9%)	2,848,353 (19.7%)		37%
40 to 64 yrs	11,201 (40.5%)	4,947,868 (34.2%)		13%
Seniors (65+ yrs)	2,050 (7.4%)	2,694,488 (18.6%)		34%
Male, n (%)	18,783 (67.9%)	7,119,636 (49.2%)	<.001	39%
Immigrated within past 10 yrs, n (%)	486 (1.8%)	624,496 (4.3%)	<.001	15%
Immigrated as refugee, n (%)	205 (0.7%)	106,988 (0.7%)	0.965	0%
Income quintile[2]				
Quintile 1 (lowest)	11,996 (43.4%)	2,722,525 (18.8%)	<.001	55%
Quintile 2	6,025 (21.8%)	2,789,370 (19.3%)		6%
Quintile 3	4,225 (15.3%)	2,894,425 (20.0%)		12%
Quintile 4	2,589 (9.4%)	2,872,394 (19.8%)		30%
Quintile 5 (highest)	2,223 (8.0%)	2,749,711 (19.0%)		32%
Missing/unknown	613 (2.2%)	456,627 (3.2%)		6%
ON-Marg subscale, N (%)[3]				
Material Deprivation				
Quintile 1 (lowest)	3,084 (11.1%)	3,173,584 (21.9%)	<.001	29%
Quintile 2	3,382 (12.2%)	2,914,508 (20.1%)		22%
Quintile 3	5,441 (19.7%)	3,280,392 (22.6%)		7%
Quintile 4	5,025 (18.2%)	2,542,710 (17.6%)		2%
Quintile 5 (highest)	10,739 (38.8%)	2,573,858 (17.8%)		48%
Dependency				
Quintile 1 (lowest)	6,161 (22.3%)	3,930,384 (27.1%)	<.001	11%
Quintile 2	6,501 (23.5%)	2,888,122 (19.9%)		9%
Quintile 3	5,631 (20.3%)	3,019,590 (20.8%)		1%
Quintile 4	4,324 (15.6%)	2,347,001 (16.2%)		2%
Quintile 5 (highest)	5,054 (18.3%)	2,299,955 (15.9%)		6%
Instability				

Quintile 1 (lowest)	1,587 (5.7%)	3,070,265 (21.2%)		47%
Quintile 2	2,064 (7.5%)	2,743,993 (18.9%)		34%
Quintile 3	4,671 (16.9%)	3,040,945 (21.0%)	<.001	11%
Quintile 4	5,312 (19.2%)	2,434,492 (16.8%)		6%
Quintile 5 (highest)	14,037 (50.7%)	3,195,357 (22.1%)		62%
Ethnic Concentration				
Quintile 1 (lowest)	3,254 (11.8%)	2,080,508 (14.4%)		8%
Quintile 2	4,158 (15.0%)	2,214,517 (15.3%)		1%
Quintile 3	6,993 (25.3%)	2,970,933 (20.5%)	<.001	11%
Quintile 4	6,164 (22.3%)	3,012,228 (20.8%)		4%
Quintile 5 (highest)	7,102 (25.7%)	4,206,866 (29.0%)		8%
Level of urbanicity, N (%)				
Large CMA (>500K)	16,110 (58.2%)	8,796,949 (60.7%)		5%
Small/Medium CMA (0-500K)	9,438 (34.1%)	3,957,887 (27.3%)		15%
Non-CMA high MIZ	579 (2.1%)	607,474 (4.2%)	<.001	12%
Non-CMA moderate/low MIZ	942 (3.4%)	672,359 (4.6%)		6%
Unknown/Missing	602 (2.2%)	450,383 (3.1%)		6%
Asthma, n (%)	4,318 (15.6%)	1,618,156 (11.2%)	<.001	13%
COPD, n (%)	3,356 (12.1%)	903,021 (6.2%)	<.001	21%
Diabetes, n (%)	3,678 (13.3%)	1,719,525 (11.9%)	<.001	4%
Congestive heart failure, n (%)	783 (2.8%)	264,975 (1.8%)	<.001	7%
Hypertension, n (%)	4,580 (16.6%)	3,073,442 (21.2%)	<.001	12%
Chronic liver disease[4], n (%)	4,054 (14.7%)	340,051 (2.3%)	<.001	45%
Rheumatoid arthritis, n (%)	221 (0.8%)	166,326 (1.1%)	<.001	4%
Charlson comorbidity index[5], n (%)				
0	6,166 (22.3%)	1,026,254 (7.1%)		44%
1	1,665 (6.0%)	160,648 (1.1%)	<.001	27%
2+	1,600 (5.8%)	218,845 (1.5%)		23%
No Hospitalizations	18,240 (65.9%)	13,079,305 (90.3%)		62%
Mental health related care[6], n (%)				
Psychotic disorders	9,134 (33.0%)	172,302 (1.2%)	<.001	93%
Non-psychotic disorders	15,884 (57.4%)	1,744,238 (12.0%)	<.001	108%
Substance use disorders	14,494 (52.4%)	205,058 (1.4%)	<.001	140%
Primary care visits[6], median (IQR)	20 (7-49)	5 (1-11)	<.001	92%

Enrolled with a primary care physician, n (%)	13,382 (48.4%)	10,887,378 (75.2%)	<.001	57%

1. Individuals at risk of homelessness. 2. Dissemination Area level income quintile, derived from census 2016 data; 3. Missing/unknown values recoded to Quintile 3; 4. Within past 3 years; 5. Within past 2 years; 6. Occurring in the past year. STD=standardized; NS=Not significant; CMA= Census Metropolitan Area; MIZ= Census metropolitan influenced zone, assigned as an estimate of the degree of influence (strong, moderate, weak or no influence) that CMAs have on the area; COPD=Chronic obstructive pulmonary disease; ON-Marg=Ontario Marginalization Index; IQR = Interquartile range

Cells representing <=5 individuals are suppressed to protect individual privacy. Immigration status defined based on presence of a landing date in the Immigration, Refugees and Citizenship Canada Permanent Resident Database from 2008 to 2018. NR = Not reportable, due to associated small cell suppression;

Table 2 - Outcomes during follow-up

Outcome	Number (% of cohort)	Rate (%) among those with a COVID-19 test	Rate (%) among those with a positive COVID-19 test	Unadjusted HR[1] (95% CI) using total cohort denominator	Unadjusted HR[1] (95% CI) Within-tested	Adjusted HR[1] (95% CI) Using total cohort denominator	Adjusted HR[1] (95% CI) Within-tested
Tested for COVID-19							
IARH[2]	3232 (11.68%)	N/A	N/A	5.32 (5.14-5.51) p<0.001	N/A	3.88 (3.74-4.02) <0.001	N/A
Community comparator	337,448 (2.33%)	N/A	N/A	-	N/A	-	N/A
Tested positive for COVID-19							
IARH[2]	180 (0.65%)	5.57%	N/A	5.01 (4.33-5.81) p< 0.001	1.09 (0.94-1.26) p=0.09	3.86 (3.33-4.49) p<0.001	1.14 (0.98-1.32) p=0.09
Community comparator	19,701 (0.14%)	5.84%	N/A	-	-	-	-
Hospitalizations with COVID-19							
IARH[2]	54 (0.20%)	1.67%	30.00%	17.82 (13.8-22.99) p<0.001	N/A	14.41 (10.99-18.90) p<0.001	N/A
Community comparator	1,702 (0.01%)	0.50%	8.64%	-	N/A	-	N/A
Death within 21 days of positive COVID-19 test							
IARH[2]	<=5	NR	NR	7.67 (3.17-18.55) p<0.001	N/A	10.79 (4.35-26.76) p<0.001	N/A
Community comparator	NR	NR	NR	-	N/A	-	N/A

1. Cox proportional hazards model. Adjusted models accounted for age, sex, immigration status, neighbourhood-level income quintile, Ontario Marginalization Index deprivation and instability subscales, level of urbanicity class, presence of comorbidities, number of outpatient visits in past year (as an indicator of healthcare usage) and enrolment with a primary care provider. 2. Individuals at risk of homelessness

Figure 1. Cohort Build

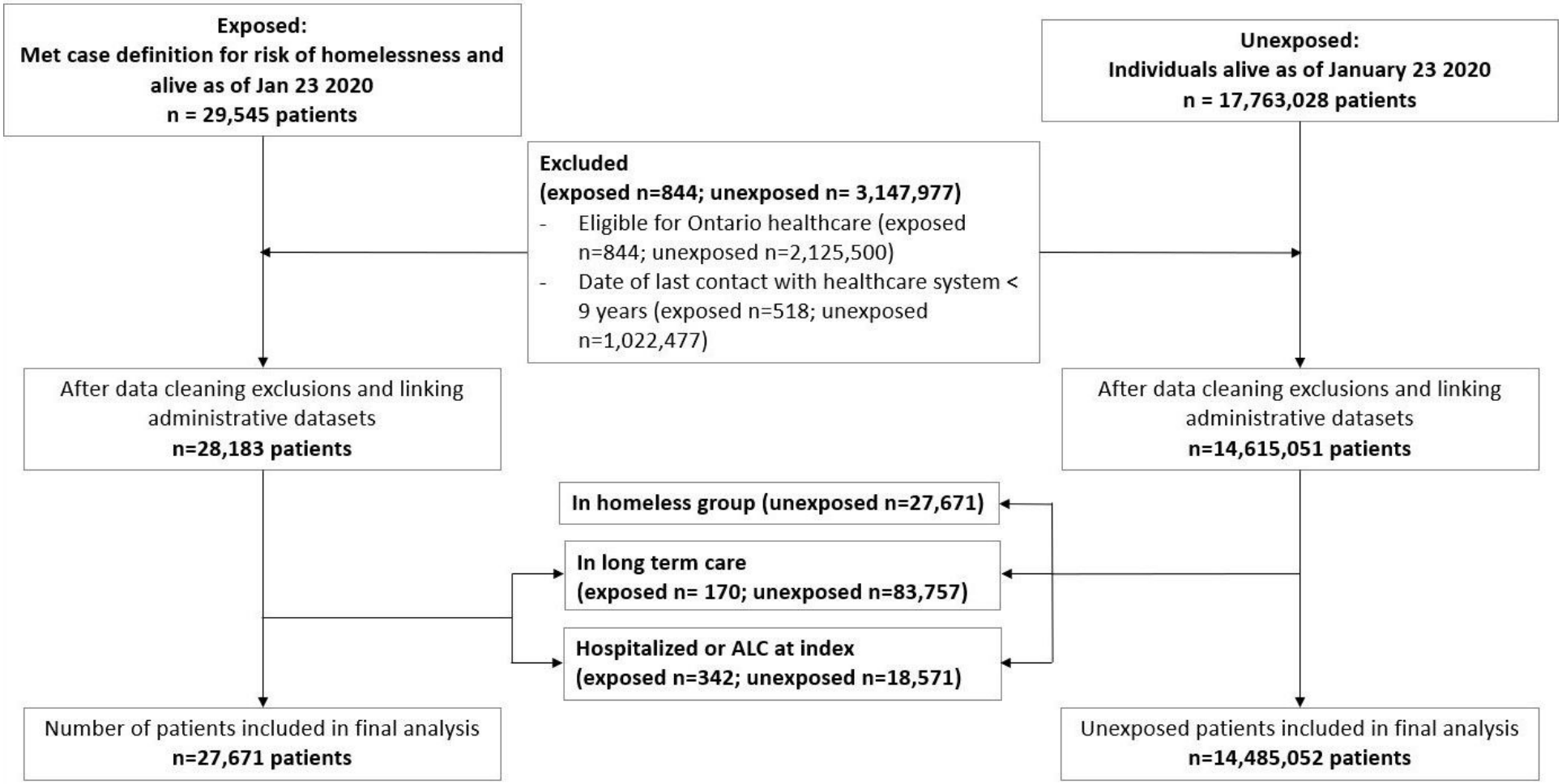


Figure 2a. COVID-19 testing rate from January 23 to May 31 2020 (per 100 persons; y-axis) among individuals identified as being at risk of homelessness per Local Health Integration Network (LHIN), ordered by decreasing population rate of individuals at risk of homelessness (x-axis)

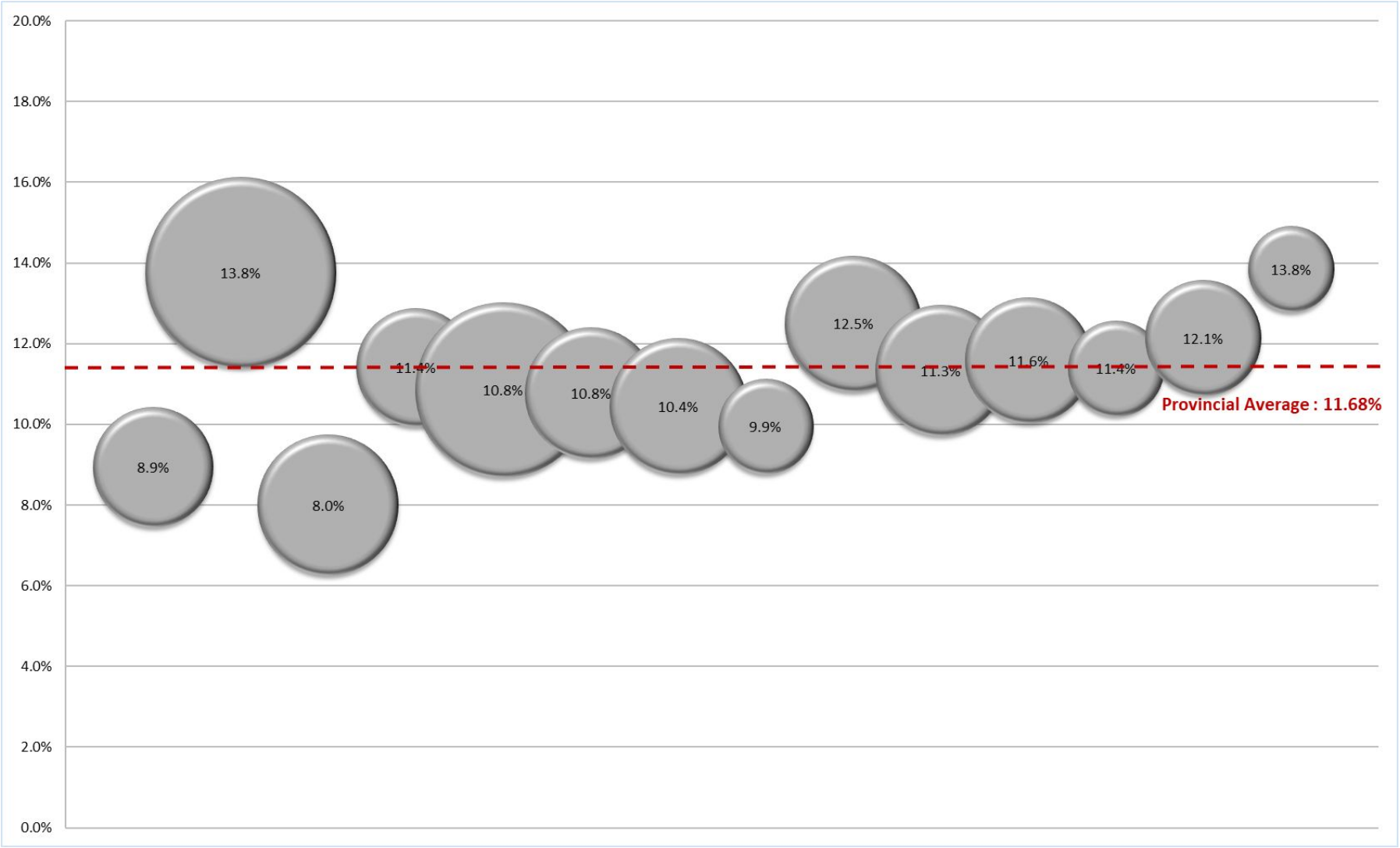


Figure 2b. COVID-19 positive rate from January 23 to May 31 2020 (per 100 persons; y-axis) among individuals at risk of homelessness who received a COVID-19 test per Local Health Integration Network (LHIN), ordered by decreasing population rate of individuals at risk of homelessness (x-axis).

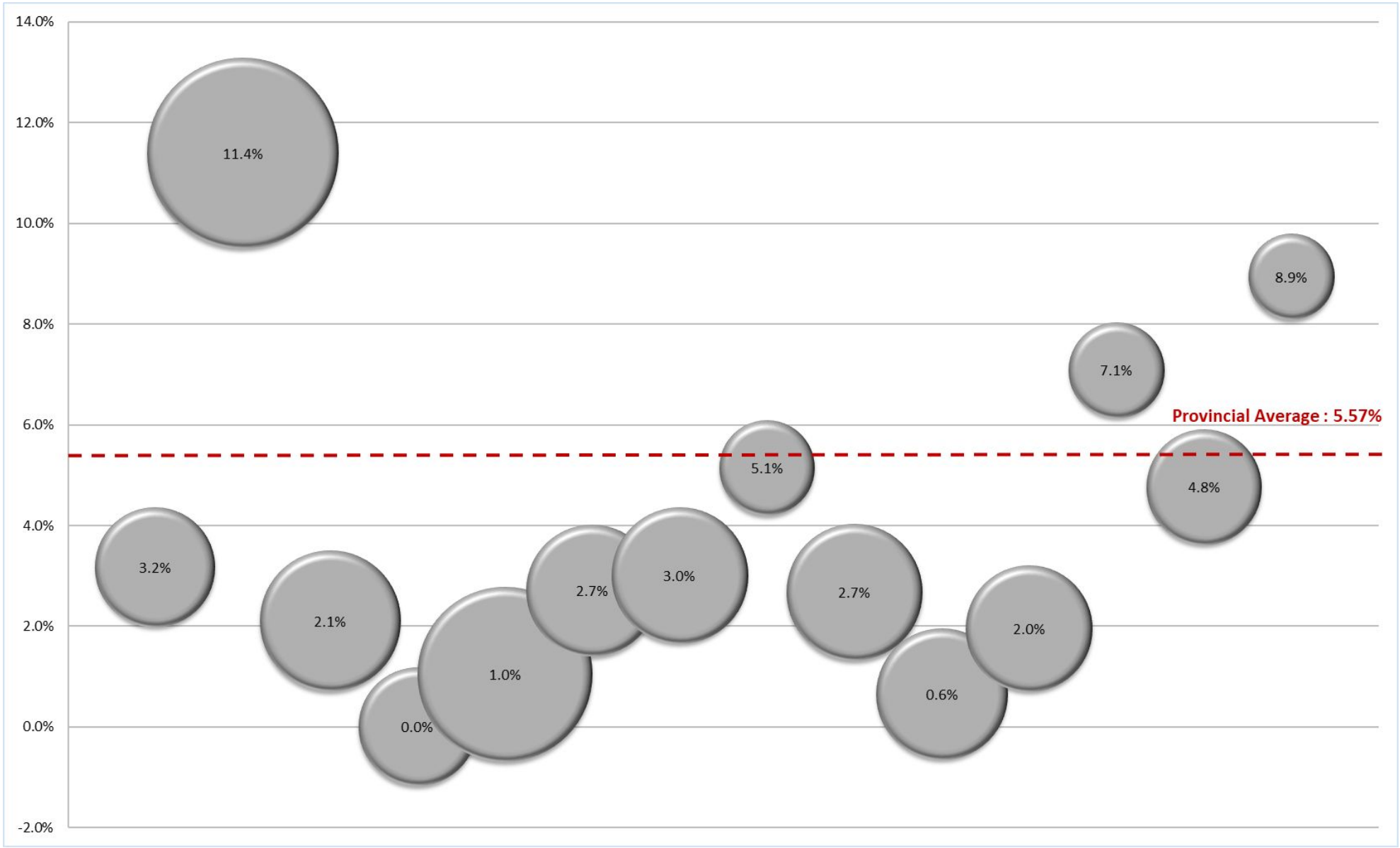
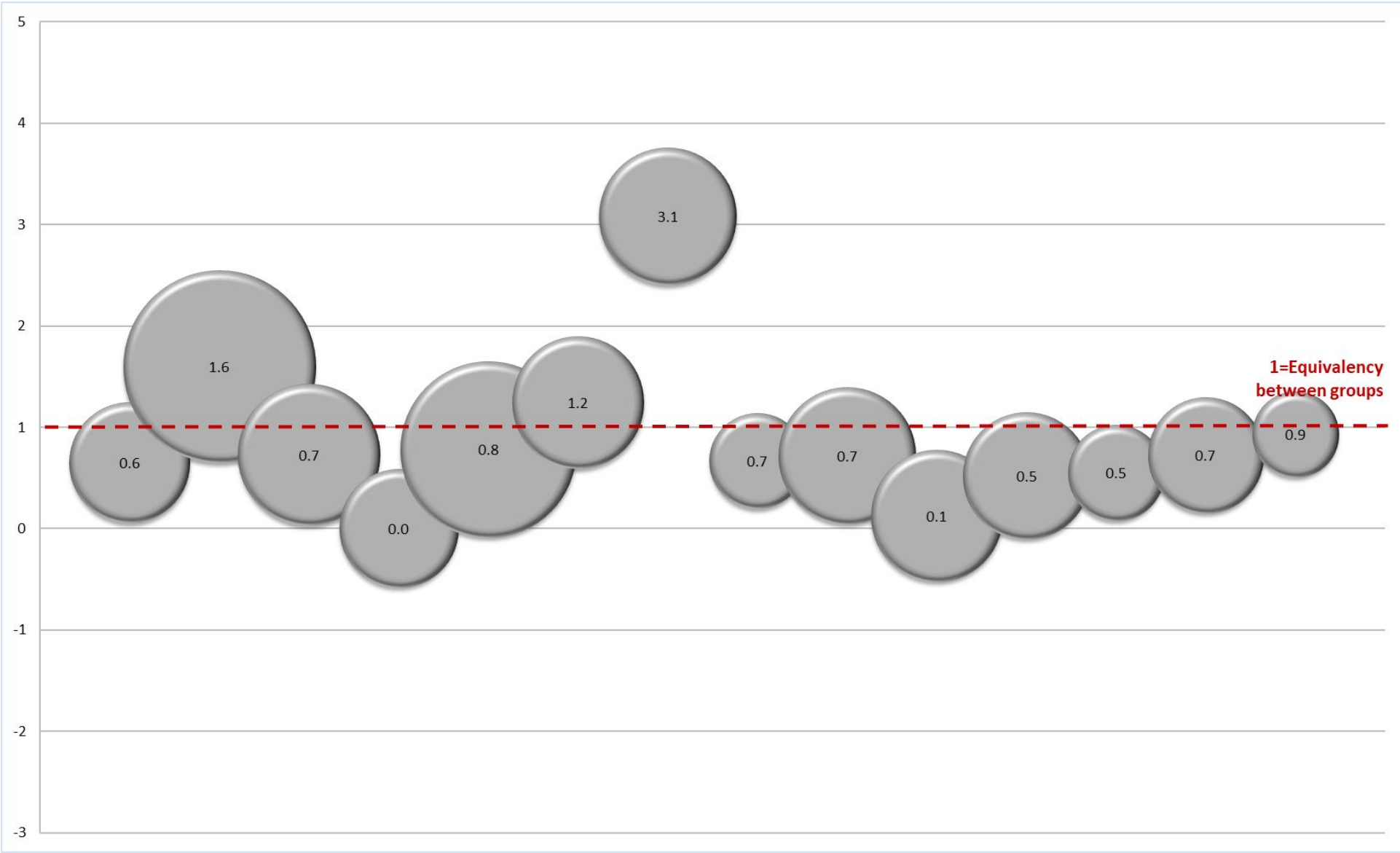


Figure 2c. Ratio of COVID-19 positive rate among individuals at risk of homelessness (who received a COVID-19 test) to COVID-19 positive rate among community-dwelling general population (who received a COVID-19 test) per Local Health Integration Network (LHIN), ordered by decreasing population rate of individuals at risk of homelessness (x-axis).



Supplement Table 1 – Case Definition of Homelessness

Case Definition = Any positive ("homeless") indicator in any of the following sources between October 1 2018 and May 31 2020:

Database	Variable Name	Indicator Value	Description
DAD	HOMELESS	"Y"	Homelessness indicator
	INSTTYPE	"SH"	Institution Type = Supportive Housing
	DX10CODE1 to DX10CODE25	"Z590" or "Z591"	ICD-10 diagnosis codes for "Homelessness" and "Inadequate housing"
	CMGDIAG	"Z590" or "Z591"	ICD-10 diagnosis codes for "Homelessness" and "Inadequate housing"
	PSTLCODE	"XX"; uniquely identifying shelter code[1]	Used to indicate transient/homeless patients
NACRS	DX10CODE1 to DX10CODE10	"Z590" or "Z591"	ICD-10 diagnosis codes for "Homelessness" and "Inadequate housing"
	RESTYPE	"3" or "4"	Residence Type = "Homeless" or "Shelter"
	PSTLCODE	"XX"; uniquely identifying shelter code[1]	Used to indicate transient/homeless patients
OMHRS	PREDX10CODE to PREDX10CODE11	"Z590" or "Z591"	ICD-10 diagnosis codes for "Homelessness" and "Inadequate housing"
	POSTDX10CODE1 to POSTDX10CODE24	"Z590" or "Z591"	ICD-10 diagnosis codes for "Homelessness" and "Inadequate housing"
	PRIOR_RESIDENCE	"6"	Prior residential status = "Homeless (with or without shelter)"
	USUAL_RESIDENCE	"8"	Usual residential status = "Homeless (with or without shelter)"
	ADMITFROM	"8"	Admitted from = "Homeless (with or without shelter)"
	DISCHLIVING	"8"	Living arrangement at discharge = "Homeless (with or without shelter)"
	P5_Retired_2009	"6"	(Variable retired in 2009) Living arrangement = "Homeless (with or without shelter)"
	PSTLCODE	"XX"; uniquely identifying shelter code[1]	Used to indicate transient/homeless patients
RPDB PSTLYEAR	PSTLCODE	uniquely identifying shelter code[1]	Used to indicate transient/homeless patients

1. ICD=International Classification of Diseases. 1. The most updated list of uniquely identifying Ontario shelter postal codes (ie. postal codes containing a homeless shelter or shelter-providing facility and no other residences) is available in Richard L, Ouédraogo AM, Shariff SZ. Identifying homelessness using administrative data and postal codes. London, ON: ICES Western; 2020.

Supplement Table 2: Databases Used

Name	Data Source	Description
Canadian Institute for Health Information Discharge Abstract Database (DAD)	Canadian Institute for Health Information (CIHI)	The DAD contains administrative, clinical (diagnoses and procedures/interventions), demographic, and administrative information for all admissions to acute care hospitals in Ontario. At ICES, consecutive DAD records are linked together to form 'episodes of care' among the hospitals to which patients have been transferred after their initial admission
Canadian Institute for Health Information Same Day Surgery (SDS)	Canadian Institute for Health Information (CIHI)	The SDS contains patient-level data admitted for same-day surgery at acute care hospitals in Ontario. Every record corresponds to one same-day surgery or procedure stay.
National Ambulatory Care Reporting System (NACRS)	Canadian Institute for Health Information (CIHI)	The NACRS contains administrative, clinical (diagnoses and procedures), demographic, and administrative information for all patient visits made to hospital- and community-based ambulatory care centres (emergency departments, day surgery units, hemodialysis units, and cancer care clinics) in Ontario. At ICES, NACRS records are linked with other data sources (DAD, Ontario Mental Health Reporting System [OMHRS]) to identify transitions to other care settings, such as inpatient acute care or psychiatric care.
Ontario Mental Health Reporting System (OMHRS)	Canadian Institute for Health Information (CIHI)	The OMHRS contains administrative, clinical (diagnoses and procedures), demographic, and administrative information for all admissions to adult designated inpatient mental health beds. This includes beds in general hospitals, provincial psychiatric facilities, and specialty psychiatric facilities. Clinical assessment data is ascertained using the Resident Assessment Instrument for Mental Health (RAI-MH), but different amounts of information are collected using this instrument depending on the length of stay in the mental health bed. Multiple assessments may occur during the length of a mental health admission.
ICES-derived PSTLYEAR database	ICES; Ministry of Health	The ICES-derived PSTLYEAR database contains the best known postal code for persons in the OHIP Registered Persons Database on July 1st of each year starting from year 1991. Postal codes supplied by the Ministry of Health are enriched with information in CIHI and other ICES-housed datasets to take advantage of the postal code information recorded each time an individual accesses certain healthcare services.

Name	Data Source	Description
Ontario Laboratory Information System (OLIS)	Ministry of Health eHealth Ontario	The OLIS contains lab orders, test requests and lab results from most laboratories in Ontario. Starting April 7 2020, ICES began receiving daily cumulative updates of COVID-19 test orders from eHealth Ontario (eHO, now part of Ontario Health Digital Health Services). These data are a minimum dataset extracted from lab orders with COVID-19-specific test request (TR) or LOINC codes and other TR/LOINC codes indicative of viral or respiratory virus testing.
OHIP Registered Persons Database	Ministry of Health	The OHIP RPDB provides basic demographic information (age, sex, location of residence, date of birth, and date of death for deceased individuals) for those issued an Ontario health insurance number. The RPDB also indicates the time periods for which an individual was eligible to receive publicly funded health insurance benefits and provides the best known postal code for each registrant on July 1st of each year.
Ontario Health Insurance Plan (OHIP)	Ministry of Health	The OHIP claims database contains information on inpatient and outpatient services provided to Ontario residents eligible for the province's publicly funded health insurance system by fee-for-service health care practitioners (primarily physicians) and "shadow billings" for those paid through non-fee-for-service payment plans. Billing codes on the claims (OHIP fee codes) identify the care provider, their area of specialization and the type and location of service. OHIP billing claims also contain a 3-digit diagnosis code - the main reason for the service - captured using a modified version of the ICD, 8th revision coding system.
Immigration, Refugees, and Citizenship Canada's Permanent Resident database (IRCC)	Immigration, Refugees and Citizenship Canada	The Ontario portion of the IRCC Permanent Resident Database includes immigration application records for people who initially applied to land in Ontario since 1985. The dataset contains permanent residents' demographic information such as country of citizenship, level of education, mother tongue, and landing date. New immigrants who are currently residing in Ontario but originally landed in another province are not captured in this dataset.
Ontario Drug Benefit (ODB) database	Ministry of Health	The Ontario Drug Benefit (ODB) database contains claims for prescription drugs received under the Ontario Drug Benefit program. Most participants of this program are over 65 years but a small number from 1997 onward are participants in other eligible programs, such as Ontario Works or the Ontario Disability Support Program.
PCCF+	Statistics Canada	The Postal Code Conversion File plus (PCCF+) provides a crosswalk between Canada Post six-character postal codes and Statistics Canada's standard geographic areas. Through the link between postal codes and standard geographic areas, the PCCF permits the integration of data from various sources. It also permits the calculation of socioeconomic status proxies, such as neighbourhood-level income quintiles.

Name	Data Source	Description
Ontario Marginalization Index	Statistics Canada; ICES	The Ontario Marginalization Index (ON-MARG) is a geographically (Census) based index developed to quantify the degree of marginalization occurring across the province of Ontario. It is comprised of 4 subscales thought to underlie the construct of marginalization: residential instability, material deprivation, dependency and ethnic concentration. The dataset at ICES applies algorithms detailed in Matheson et al. "Development of the Canadian Marginalization Index: a new tool for the study of inequality." Canadian Journal of Public Health, 2012;103(Suppl. 2):S12-S16.
CAPE	Ministry of Health	The Client Agency Program Enrolment (CAPE) contains patient enrolment with primary care physicians in Ontario. Individuals must be eligible for health coverage to be included.
Chronic Obstructive Pulmonary Disease (COPD) dataset	Canadian Institute for Health Information (CIHI)	The Ontario COPD Database is created using two separate algorithms applied to inpatient hospitalization (DAD), same day surgery (SDS) records, and physician billing claims (OHIP) data to determine the diagnosis date for incident cases of chronic obstructive pulmonary disease in Ontario. In an algorithm which maximizes sensitivity, the definition for COPD is any physician billing claim with a diagnosis for COPD (OHIP diagnosis codes: 491, 492, 496) or any inpatient hospitalization or same day surgery record with a diagnosis for COPD (ICD-9 diagnosis codes: 491, 492, 496; ICD-10 diagnosis codes: J41- J44; in any diagnostic code space). When using expert panel review of primary care charts as the reference standard, this definition has been shown to have the following performance characteristics: Sensitivity (85.0%), Specificity (78.4%), Positive Predictive Value (57.5%), and Negative Predictive Value (93.8%).(7) In an algorithm which maximizes specificity, the definition for COPD is ≥3 physician billing claims with a diagnosis for COPD (OHIP diagnosis codes: 491, 492, 496) or ≥1 inpatient hospitalization or same day surgery record with a diagnosis for COPD (ICD-9 diagnosis codes: 491, 492, 496; ICD-10 diagnosis codes: J41, J42, J43, J44; in any diagnostic code space) in a two-year period. When using expert panel review of primary care charts as the reference standard, this definition has been shown to have the following performance characteristics: Sensitivity (57.5%), Specificity (95.4%), Positive Predictive Value (81.3%), and Negative Predictive Value (86.7%).(1)

Name	Data Source	Description
Ontario Asthma dataset	Canadian Institute for Health Information (CIHI)	The Ontario Asthma Database is created using two separate algorithms applied to inpatient hospitalization (DAD), same day surgery (SDS) records, and physician billing claims (OHIP) data to determine the diagnosis date for incident cases of asthma in Ontario. In the algorithm which maximized sensitivity, the definition for Asthma is receipt of one hospital admission with an asthma diagnosis or two OHIP claims with asthma diagnosis within two years. This definition has been shown to have the following performance characteristics in adults: Sensitivity (80.6%), Specificity (81.4%), Positive Predictive Value (72.5%), and Negative Predictive Value (87.3%).(2)
Ontario Diabetes Database (ODD)	Canadian Institute for Health Information (CIHI)	The ODD is created using algorithms applied to inpatient hospitalization (DAD) records, same day surgery (SDS) records, and physician billing claims (OHIP) data to determine the diagnosis date for incident cases of diabetes in Ontario. For adults aged 19 years and greater, the definition for diabetes is 2 physician billing claims with a diagnosis for diabetes (OHIP diagnosis code: 250) or 1 inpatient hospitalization or same day surgery record with a diagnosis for diabetes (ICD-9 diagnosis code: 250; ICD-10 diagnosis codes: E10, E11, E13, E14; in any diagnostic code space) within a 2 year period. Physician claims and hospitalizations with a diagnosis of diabetes occurring within 120 prior to and 180 days after a gestational hospitalization record were excluded. When using primary care chart abstraction as the reference standard, this definition has been shown to have the following performance characteristics: Sensitivity (86.1%), Specificity (97.1%), Positive Predictive Value (79.8%), and Negative Predictive Value (98.1%).(3)
Ontario CHF Database (CHF)	Canadian Institute for Health Information (CIHI)	The Ontario CHF Database is created using a definition of ≥2 physician billing claims with a diagnosis of congestive heart failure (OHIP diagnosis code: 428) and/or ≥1 inpatient hospitalization or same day surgery record with a diagnosis of congestive heart failure (ICD-9 diagnosis code: 428; ICD-10 diagnosis code: I50; in the primary diagnostic code space) in a two-year period applied to hospitalization (DAD), same day surgery (SDS), and physician billing claims (OHIP) data to determine the diagnosis date for incident cases of CHF in Ontario. When using electronic medical record data abstraction as the reference standard, the above definition has been demonstrated to have the following performance characteristics: Sensitivity (84.8%), Specificity (97.0%), and Positive Predictive Value (55.3%).(4)

Name	Data Source	Description
Ontario Hypertension dataset	ICES	The Ontario hypertension Dataset contains all Ontario hypertension patients identified since 1991. The case definition is one hospital admission/SDS record with a hypertension diagnosis (ICD 9 dxcodes: 401x, 402x, 403x, 404x, 405x or ICD 10 dx10codes: I10, I11, I12, I13, I15), or 1 OHIP claim (401, 402, 403, 404, or 405) with a hypertension diagnosis followed by either an OHIP claim or a hospital admission/SDS record with a hypertension diagnosis within two years. If the hypertension record is between 120 days before and 180 days after a gestational admission date, the hypertension record was considered to be gestational hypertension, and was excluded. This definition has been shown to have the following performance characteristics in adults: Sensitivity (72%), Specificity (95%), Positive Predictive Value (87%)(5)
Ontario Rheumatoid Arthritis dataset	ICES	The Ontario Rheumatoid Arthritis Database (ORAD) contains all Ontario rheumatoid arthritis patients identified since 1991. The case definition is receipt of at least 1 inpatient CIHI DAD diagnosis code (any type) for RA (ICD9 714; ICD10 M05, M06) or at least 3 OHIP claims with a diagnosis code for RA (dx 714) over a 2 year period, with at least 1 claim by a musculoskeletal specialist. This definition has been shown to have the following performance characteristics in adults: Sensitivity (78%), Specificity (100%), Positive Predictive Value (78%) and Negative Predictive Value (100%)(6)

Supplement Table 3: Variable Definitions

Variable	Data Source	Definition Description
Age	RPDB	Age of the individual at index
Sex	RPDB	Biological sex of the individual
Neighbourhood level income quintile	Census	Income-level assigned to the individual's dissemination area at index, expressed as a quintile (compared to all other dissemination areas that year)
Ontario Marginalization Index	ONmarg	Ontario Marginalization Index (ON-MARG) is a geographically (Census) based index developed to quantify the degree of marginalization occurring across the province of Ontario. It is comprised of 4 major dimensions thought to underlie the construct of marginalization: residential instability, material deprivation, dependency and ethnic concentration. Expressed as a quintile, in comparison to all other dissemination areas that year.
Level of urbanicity	PCCF	Adapted from CSIZEMIZ variable which categorizes individuals based on size of census metropolitan area (CSIZE) and area's degree of influence by a nearby census metropolitan area (MIZ). Medium and small CMAs (0-500K pop) are grouped together, as are non-CMA areas with moderate or low MIZ.
Recent immigrant	IRCC	Presence of a landing date in the Immigration, Refugees and Citizenship Canada Permanent Database indicates immigration to Ontario between 2008 to 2018
Refugee status	IRCC	Class of immigration status = Refugee
Local Health Integration Network (LHIN)	RPDB	Local Health Integration Networks are Ontario's health authorities responsible for administration of public healthcare. This variable contains the LHIN in which the individual is believed to reside as of the index, based on their census division information
Charlson comorbidity index	DAD	
Chronic obstructive pulmonary disease	COPD	Presence in the database indicates the individual has a history of COPD
Asthma	ASTHMA	Presence in the database indicates the individual has a history of ASTHMA
Diabetes	ODD	Presence in the database indicates the individual has a history of diabetes
Congestive heart failure	CHF	Presence in the database indicates the individual has a history of congestive heart failure
Hypertension	HYPER	Presence in the database indicates the individual has a history of hypertension

Variable	Data Source	Definition Description
Chronic liver disease	DAD, NACRS, OHIP	1 hospitalization, ED visit or physician claim within 3 years of the index date with any of the following eligible codes: ICD-10: B16, B17, B18, B19, B942, E830, E831, I85, K70, K713, K714, K715, K717, K721, K729, K73, K74, K753, K754, K758, K759, K76, K77, R160, R162, R17, R18, Z225 OHIP dx: 070, 571, 573 OHIP fee: Z551, Z554
Rheumatoid Arthritis	ORAD	Presence in the database indicates the individual has a history of rheumatoid arthritis
Psychosis related mental health care	DAD, NACRS, OMHRS, OHIP	1 hospitalization, ED visit or physician claim within 1 year of the index date with any of the following eligible codes: ICD-10: F20, F22, F23, F24, F25, F28, F29 DSM-IV: 295, 297, 298 OHIP dx: 295, 297, 298
Non-psychotic disorders related mental health care	DAD, NACRS, OMHRS, OHIP	1 hospitalization, ED visit or physician claim within 1 year of the index date with any of the following eligible codes: ICD-10: F30, F31, F32, F33, F34, F38, F39, F40, F41, F42, F43, F48, F60, F93 DSM-IV: 296, 300, 301 OHIP dx: 296, 300, 301, 309, 311
Substance use related mental health care	DAD, NACRS, OMHRS, OHIP	1 hospitalization, ED visit or physician claim within 1 year of the index date with any of the following eligible codes: ICD-10: F10, F11, F12, F13, F14, F15, F16, F17, F18, F19, F55 DSM-IV: 291, 292, 303, 304, 305 OHIP dx: 291, 292, 303, 304, 305
Outpatient visits	OHIP	Number of physician visits within 1 year prior to the index date, defined as one visit per day per physician
Enrolment with a family physician	CAPE	

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	A) Title/abstract B) Abstract methods and results	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	1.1) Abstract (methods) 1.2) Title; Abstract (methods) 1.3) Abstract (methods)
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction, par 1		
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction, par 2		
Methods					
Study Design	4	Present key elements of study design early in the paper	Methods, "Study Design and Setting"		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Methods, "Study Design and Setting"		

Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case	A) Methods, "Participants" B) No matching criteria (N/A)	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	6.1) Provided in Supplement Tables 6.2) Validations are referenced
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	Methods "Outcome Measures" and "Other Covariates"	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	7.1) Provided in Supplement Tables
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Methods "Data Sources" and Supplement Tables		

Bias	9	Describe any efforts to address potential sources of bias	N/A		
Study size	10	Explain how the study size was arrived at	N/A		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	Methods "Other Covariates"		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	Methods "Analysis"		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.	12.1) Methods "Study Design and Setting" 12.2) N/A (ICES study)

				RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	12.3) Methods "Study Design and Setting"
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	Results; Figure 1	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	13.1) Results; Fig 1
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)	Results; Table 1		
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure	Results; Table 2		

		category, or summary measures of exposure *Cross-sectional study* - Report numbers of outcome events or summary measures			
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Results; Table 2		
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	N/A		
Discussion					
Key results	18	Summarise key results with reference to study objectives	Discussion, par 1&2		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Discussion, par 3	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Discussion, limitations paragraph
Interpretation	20	Give a cautious overall interpretation of results considering objectives,	Discussion, overall		

		limitations, multiplicity of analyses, results from similar studies, and other relevant evidence			
Generalisability	21	Discuss the generalisability (external validity) of the study results	Discussion, par 3		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Acknowledgements; Title page		
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	N/A (ICES study)

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.