

# Utility, Limitations and Opportunities for Using Linked Health Administrative Data to Study Homelessness in Ontario

Richard G. Booth\*, Lucie Richard\*, Cheryl Forchuk and Salimah Z. Shariff

## Abstract

**Homelessness is a significant social issue within Canada but is difficult to quantify at the population level. In this paper, we discuss the development and use of a case ascertainment algorithm that identifies people experiencing homelessness through health administrative data. We highlight the appropriateness of various uses of this method given its key strengths and limitations. Finally, we discuss plans to improve this methodology and broaden its use through the addition of linkable administrative data from non-health sectors, such as emergency shelters and social services organizations.**

## Introduction

Homelessness is a significant social and healthcare issue. People experiencing homelessness (PEH) face significant social stigma and numerous structural inequities and are at increased risk for a variety of poor health outcomes (Bowen et al. 2019). This population is highly heterogeneous, with many subgroups, such as Indigenous People and veterans, who are disproportionately represented; these subgroups experience context-dependent challenges that require specific types and intensities of support (Findlay et al. 2018; Kidd et al. 2017; Kneebone et al. 2015; Pomeroy 2020; Schiff et al. 2016; Strobel et al. 2021). Despite this, our ability to respond effectively to homelessness is hindered by difficulties in counting PEH (CAEH 2018). In Canada, the most common approach used is the point-in-time (or PiT) count, which many communities deploy to ascertain the minimum number of individuals experiencing homelessness locally on a given night (Echenberg and Munn-Rivard 2020). During this event, efforts are made to collect demographic information and housing history. Many communities have also implemented information management systems, such as

the Homeless Individuals and Families Information System (HIFIS), to harmonize data gathering at homeless-serving organizations, including emergency shelters. It is from these approaches that Gaetz et al. (2016) have estimated that approximately 235,000 individuals experience homelessness in a given year in Canada.

Unfortunately, these measurement approaches leave important gaps in our understanding of homelessness in Canada. PiT counts are cross-sectional in nature, conducted irregularly and extremely resource intensive (Echenberg and Munn-Rivard 2020) and are only able to identify individuals who are visibly homeless or living in shelters. As such, these counting methods can miss *hidden* homeless populations, including those who are precariously housed, living in the rough or couch surfing. Furthermore, traditional PiT approaches can be difficult to derive into representative community-based estimates (Williams 2011) and do not permit individual-level follow-up over time. To the extent that HIFIS and similar information management systems cannot link records to an individual – whether because informed consent or reliable information to conduct linkage (or both) is not available – they have the same limitation. As a result, it is not possible to use current methods to evaluate the impact of programs and policies aimed at relief from and prevention of homelessness without setting aside funds from these programs for the collection of data for evaluation purposes or to assess changes in the size of the population over time. This is particularly true during public health crises or other events that hinder data collection, such as the COVID-19 pandemic, where PiT counts were largely postponed (Echenberg and Munn-Rivard 2020).

Given all of the above-mentioned points, there is value in developing a cost-effective and scalable measurement method

\*Co-primary authors

for PEH that permits follow-up over time and is less sensitive to crises that prevent deliberate measurement. Here, we briefly describe a case ascertainment algorithm that the authors developed in 2019 to identify PEH using health administrative data. We highlight the current utility and limitations of this algorithm for research and evaluation and list future developments and opportunities in the use of administrative data for research and evaluation in service of PEH across Canada.

### Using Health Administrative Data to Measure Homelessness

In 2019, we validated a case ascertainment algorithm that uses health administrative data, which is information routinely collected during healthcare encounters (Richard et al. 2019). To do this, we compared a group of individuals whose housing status had been recorded over several years by Hwang et al. (2011) with health administrative data available from data repositories such as ICES ([www.ices.on.ca](http://www.ices.on.ca)) to determine if the administrative data could be used to correctly and consistently identify PEH. We found that these data identified a portion of PEH (with a sensitivity of 10–35% depending on the algorithm used) but almost never incorrectly identified someone as experiencing homelessness if they were not (with a specificity exceeding 99%).

Since the development of this algorithm, various studies have leveraged this method to measure or describe homelessness in Ontario and Canada (Hossain et al. 2021; Richard et al. 2021; Strobel et al. 2021; Wang et al. 2020). There are obvious advantages to the use of this tool, which explains its immediate uptake. First, it is relatively inexpensive and readily leveraged to gain population-level insights that can inform policy and practice. This was particularly valuable at the beginning of the COVID-19 pandemic, when the emergent nature of the situation required novel and expedited measurement methods to inform public health protocols. Furthermore, data housed at repositories, such as ICES, are fully de-identified and heavily protected by policies and procedures that ensure the privacy and confidentiality of individuals. Finally, health administrative data are routinely linked, permitting passive follow-up of identified persons through healthcare encounters over time; this is difficult or impossible to conduct with other measurement methods.

The optimal use of this algorithm, in its current form, is to create convenience samples of PEH for longitudinal research and evaluation. With some care, the algorithm can also be used to compare the risk of a given health outcome or healthcare utilization pattern among PEH with that of a housed population and thereby highlight excess risks associated with being homeless. For example, we recently used the algorithm to compare estimates of COVID-19 infection and complication rates among PEH in Ontario to those of the

general, community-dwelling population (Richard et al. 2021). We are also currently assessing the impact of the COVID-19 pandemic on patterns of healthcare utilization, vaccine uptake and vaccine efficacy among people recently made homeless. In these projects, our questions could only be adequately and quickly answered with a reasonably reliable denominator and the ability to follow identified persons over time, something only this method currently permits on a population level.

Despite our methodology's advantages, several important limitations must be acknowledged. Firstly, data are linked at ICES and similar institutions using an identifier or a set of identifiers – in the case of ICES, the identifier is derived from the Ontario Health Insurance Plan cards. As a result, those without provincial health coverage are, by definition, not included in cohorts that are generated using this method. This disproportionately excludes Indigenous People, refugees and certain veterans who have alternative federal health coverage. As many of these groups are also disproportionately represented among PEH, this is a key limitation that impedes the generalizability of results.

Secondly, to be identified by this method, individuals must interact with tertiary or hospital care. Those with acute health conditions, greater healthcare-seeking behaviour or longer periods of time spent homeless are more likely to be represented in cohorts developed using this method. Conversely, healthier persons, those avoiding hospital-based care and those experiencing short periods of homelessness will be less represented. Although all research using health administrative data inherently suffers from this limitation to some extent, algorithms designed to identify a changeable social determinant of health such as homelessness are particularly hindered as the experience of homelessness does not, by itself, mandate healthcare encounters.

Finally, this method in its current form is restricted to enumerating homelessness from a rather narrow epidemiological interpretation of the phenomena. The hidden homeless (e.g., those precariously housed, couch surfers), who are so difficult to capture in all measurement methods, will likely also be systematically missed using this method, unless they choose to disclose their status during an emergency department visit or hospitalization. Moreover, health administrative data on their own provide very little social information (e.g., self-reported gender, race, language, income, marital status, housing instability and geographic and built-environment considerations) to qualify and enrich our understanding of the impact of homelessness on healthcare and health outcomes.

### Recommendations for Use

Those wishing to use this method for research or evaluation in service of PEH must assess whether their intended use is appropriate, having considered the above-mentioned advantages and

limitations. The algorithm's low sensitivity means that it is generally inappropriate to use it to characterize a subgroup as PEH within a larger population as is often sought by researchers whose question is focused on the general population. While it is understandable that one might want to assign homelessness as a proxy for social vulnerability, leveraging the algorithm in this fashion will almost always generate a significant underestimate of the true number of PEH. This can, at best, create ineffective covariate adjustment but also easily lead to biased group effect estimates and unintentional misinterpretation by researchers and policy makers.

Furthermore, this method should not be used to make inferences about PEH it cannot properly identify, particularly known subgroups such as Indigenous People, refugees, veterans, the hidden homeless and those who avoid hospital-based healthcare. When conducting research on long-term or recent experiences of homelessness generally, it is vital to highlight the extent to which results are *not* generalizable. For example, in our recent study on COVID-19 infection and complication rates (Richard et al. 2021), we indicated that our method was ineffective in identifying PEH who are Indigenous or refugees and suggested that our study results could only be generalized to individuals who accessed health services.

### Future Opportunities

The current algorithm provides a valuable tool to help identify PEH, particularly over time. However, despite the recentness of our validation work, the social and data landscapes have changed significantly since its publication. Some key identifiers of homelessness used by the algorithm (including diagnostic codes Z590 and Z591 from the *International Classification of Diseases, 10th edition* [CIHI 2018]) have since April 2018 become mandatory for abstractors to code where they are documented in patient charts (CIHI 2021), which may increase the sensitivity of capture. Furthermore, new administrative data at ICES, such as the recently linked Community Health Centre data set (Booth et al. 2020; Glazier et al. 2012), have provided new opportunities to identify PEH who do not use traditional hospital-based healthcare. Finally, emerging evidence suggests that the experience of homelessness itself has changed since the COVID-19 pandemic. There appears to be greater visibility in rural settings (Devereaux 2021; Schiff et al. 2020) where previously homelessness was either not present or hidden, more reliance on temporary hotel accommodations or urban camping in avoidance of emergency shelters or as a

result of reduction in affordable housing (Baral et al. 2021; Devereaux 2021) and potentially more reliance on hospital-based care due to reductions in or closures of homeless-serving programs and services (Gibson et al. 2020; Perri et al. 2020). If these changes persist, the algorithm may perform differently than before the pandemic. For all these reasons, revalidation of the algorithm using health administrative data collected after 2018 – and ideally in the post-pandemic period – is indicated.

In the longer term, there is tremendous opportunity to enhance identification of PEH using administrative data outside of the healthcare system. Data repositories with established, rigorous protocols for linking data securely and flexibly to individuals (without the need for non-universally available identifiers such as provincial healthcare coverage) could host administrative data from other agencies that serve PEH (e.g., emergency shelters, social services organizations and HIFIS). Triangulation between different service sectors would fill, to the extent possible, gaps inherent in most measurement methods and generate the most representative population of PEH possible. While the barriers to achieving this are notable (e.g., privacy concerns and the need for legal data-sharing agreements), these data could also provide a better understanding of the complex realities faced by PEH beyond their impact on the healthcare system, which in itself would greatly benefit research and service provision for this vulnerable population.

### Conclusion

Measuring homelessness is vital to guide an evidence-based approach to relief and prevention. Use of administrative data, such as health administrative data at ICES, can help improve identification of PEH and permit longitudinal follow-up. However, use must be appropriate and must acknowledge the key limitations of the methodology and, wherever possible, combine this method with other sources of data to enhance the method to develop specific and accurate insights. With the COVID-19 pandemic and housing crisis, this is more imperative than ever as without contemporary understanding and measurement of homelessness, solutions cannot be adequately developed and implemented to address the increasing complexity and scale of PEH in Canada. **HQ**

### Acknowledgement

Parts of the manuscript were supported by the Public Health Agency of Canada.

### References

- Baral, S., A. Bond, A. Boozary, E. Bruketa, N. Elmi, D. Freiheit et al. 2021. Seeking Shelter: Homelessness and COVID-19. *Facets* 6: 925–58. doi:10.1139/facets-2021-0004.
- Booth, R.G., L. Richard, L. Li, S.Z. Shariff and J. Rayner. 2020. Characteristics of Health Care Related to Mental Health and Substance Use Disorders among Community Health Centre Clients in Ontario: A Population-Based Cohort Study. *CMAJ Open* 8(2): E391–99. doi:10.9778/cmajo.20190089.

- Bowen, M., S. Marwick, T. Marshall, K. Saunders, S. Burwood, A. Yahyouche et al. 2019. Multimorbidity and Emergency Department Visits by a Homeless Population: A Database Study in Specialist General Practice. *British Journal of General Practice* 69(685): e515–25. doi:10.3399/bjgp19X704609.
- Canadian Alliance to End Homelessness (CAEH). 2018, May 9. Why Real-Time Data Is the Foundation for Ending Homelessness. Retrieved October 19, 2021. <<https://caeh.ca/real-time-data/>>.
- Canadian Institute for Health Information (CIHI). 2018. Version 2018 ICD-10-CA and CCI Classifications. Retrieved December 20, 2021. <<https://www.cihi.ca/en/version-2018-icd-10-ca-and-cci-classifications>>.
- Canadian Institute for Health Information (CIHI). 2021, March 22. Updated ICD-10-CA Coding Direction: Homelessness, and Falls from an Electric Scooter (Escooter), Mobility Scooter, Segway® or Hoverboard. Retrieved October 19, 2021. <<https://www.cihi.ca/en/bulletin/updated-icd-10-ca-coding-direction-homelessness-and-falls-from-an-electric-scooter>>.
- Devereaux, A. 2021, October 19. Rural Canada's 'Visible Homelessness' Problem Driven by Pandemic, High Cost of Housing. *CBC News*. Retrieved October 19, 2021. <<https://www.cbc.ca/news/canada/london/homeless-rural-ontario-huron-county-1.6204515>>.
- Echenberg, H. and L. Munn-Rivard. 2020, December 18. *Defining and Enumerating Homelessness in Canada* [Background paper]. Library of Parliament. Retrieved October 19, 2021. <<https://lop.parl.ca/staticfiles/publicwebsite/home/researchpublications/backgroundpapers/pdf/2020-41-e.pdf>>.
- Findlay, I. M., J. Chilima, B. Holden and A. Berthe. 2018. *2018 Point-in-Time Homelessness Count, Saskatoon, Saskatchewan*. Community-University Institute for Social Research. Retrieved October 19, 2021. <[https://cuirs.usask.ca/documents/publications/2015-2019/CUISR\\_PIT2018\\_JULY%204TH\\_2019.pdf](https://cuirs.usask.ca/documents/publications/2015-2019/CUISR_PIT2018_JULY%204TH_2019.pdf)>.
- Gaetz, S., E. Dej, T. Richter and M. Redman. 2016. *The State of Homelessness in Canada 2016*. COH Research Paper #12. Canadian Observatory on Homelessness. Retrieved October 19, 2021. <[https://homelesshub.ca/sites/default/files/sohc16\\_final\\_20oct2016.pdf](https://homelesshub.ca/sites/default/files/sohc16_final_20oct2016.pdf)>.
- Gibson, V. 2020, November 10. St. Michael's Hospital Develops a Tool to Track Homeless Patients — With Hopes of Detecting COVID-19 Outbreaks Early. *Toronto Star*. Retrieved October 19, 2021. <<https://www.thestar.com/news/gta/2020/11/09/st-michaels-hospital-develops-a-tool-to-track-homeless-patients-with-hopes-of-detecting-covid-19-outbreaks-early.html>>.
- Glazier, R.H., B.M. Zagorski and J. Rayner. 2012, March. *Comparison of Primary Care Models in Ontario by Demographics, Case Mix and Emergency Department Use, 2008/09 to 2009/10*. ICES Investigative Report. Institute for Clinical Evaluative Sciences. Retrieved October 19, 2021. <<https://www.ices.on.ca/-/media/Files/Atlases-Reports/2012/Comparison-of-primary-care-models-in-Ontario/Full-report.ashx>>.
- Hossain, R., J.H. Dai, S. Jamani, Z. Ma, E. Dvorani, E. Graves et al. 2021. Hard-to-Reach Populations and Administrative Health Data: A Serial Cross-Sectional Study and Application of Data to Improve Interventions for People Experiencing Homelessness. *Medical Care* 59(Suppl. 2): S139–45. doi:10.1097/mlr.0000000000001481.
- Hwang, S.W., T. Aubry, A. Palepu, S. Farrell, R. Nisenbaum, A.M. Hubley et al. 2011. The Health and Housing in Transition Study: A Longitudinal Study of the Health of Homeless and Vulnerably Housed Adults in Three Canadian Cities. *International Journal of Public Health* 56(6): 609–23. doi:10.1007/s00038-011-0283-3.
- Kidd, S.A., S. Gaetz and B. O'Grady. 2017. The 2015 National Canadian Homeless Youth Survey: Mental Health and Addiction Findings. *Canadian Journal of Psychiatry* 62(7): 493–500. doi:10.1177/0706743717702076.
- Kneebone, R.D., M. Bell, N. Jackson and A. Jadidzadeh. 2015. Who Are the Homeless? Numbers, Trends and Characteristics of Those without Homes in Calgary. *SPP Research Paper* 8(11): 1–17. doi:10.11575/sppp.v8i0.42510.
- Perri, M., N. Dosani and S.W. Hwang. 2020. COVID-19 and People Experiencing Homelessness: Challenges and Mitigation Strategies. *CMAJ* 192(26): E716–19. doi:10.1503/cmaj.200834.
- Pomeroy, S. 2020, July. *Recovery for All: Proposals to Strengthen the National Housing Strategy and End Homelessness*. Canadian Alliance to End Homelessness. Retrieved October 19, 2021. <<https://caeh.ca/wp-content/uploads/Recovery-for-All-Report-July-16-2020.pdf>>.
- Richard, L., S.T. Hwang, C. Forchuk, R. Nisenbaum, K. Clemens, K. Wiens et al. 2019. Validation Study of Health Administrative Data Algorithms to Identify Individuals Experiencing Homelessness and Estimate Population Prevalence of Homelessness in Ontario, Canada. *BMJ Open* 9:e030221. doi:10.1136/bmjopen-2019-030221.
- Richard, L., R. Booth, J. Rayner, K.K. Clemens, C. Forchuk and S.Z. Shariff. 2021. Testing, Infection and Complication Rates of COVID-19 among People with a Recent History of Homelessness in Ontario, Canada: A Retrospective Cohort Study. *CMAJ Open* 9(1): E1–E9. doi:10.9778/cmajo.20200287.
- Schiff, J.W., R. Schiff and A. Turner. 2016. Rural Homelessness in Western Canada: Lessons Learned from Diverse Communities. *Social Inclusion* 4(4): 73–85. doi:10.17645/si.v4i4.633.
- Schiff, R., K. Buccieri, J.W. Schiff, C. Kauppi and M. Riva. 2020. COVID-19 and Pandemic Planning in the Context of Rural and Remote Homelessness. *Canadian Journal of Public Health* 111(6): 967–70. doi:10.17269/s41997-020-00415-1.
- Strobel, S., I. Burcul, J.H. Dai, Z. Ma, S. Jamani and R. Hossain. 2021. Characterizing People Experiencing Homelessness and Trends in Homelessness Using Population-Level Emergency Department Visit Data in Ontario, Canada. *Health Reports* 32(1): 13–23. doi:10.25318/82-003-x202100100002-eng.
- Wang, L., H. Ma, K.C.Y. Yiu, A. Calzavara, D. Landsman, L. Luong et al. 2020. Heterogeneity in Testing, Diagnosis and Outcome in SARS-Cov-2 Infection across Outbreak Settings in the Greater Toronto Area, Canada: An Observational Study. *CMAJ Open* 8(4): E627–36. doi:10.9778/cmajo.20200213.
- Williams, J.C. 2011. "Stand Up and Be Counted": The Politics of a Homeless Enumeration. *Poverty & Public Policy* 3(3): 4. doi:10.2202/1944-2858.1172.

## About the authors

**Richard G. Booth**, PhD, RN, is an associate professor at the Arthur Labatt Family School of Nursing at Western University in London, ON. He can be contacted by e-mail at [rbooth6@uwo.ca](mailto:rbooth6@uwo.ca).

**Lucie Richard**, MA, is a health geographer and analytic epidemiologist at ICES Western in London, ON. She can be contacted by e-mail at [lucie.richard@ices.on.ca](mailto:lucie.richard@ices.on.ca).

**Cheryl Forchuk**, PhD, RN, is the Beryl and Richard Ivey Research Chair in Aging, Mental Health, Rehabilitation and Recovery, a distinguished professor in Nursing and Psychiatry at Western University and an assistant director at the Lawson Health Research Institute in London, ON.

**Salimah Z. Shariff**, PhD, is a staff scientist at ICES Western, an associate scientist at the Lawson Health Research Institute and an adjunct research professor in the Arthur Labatt Family School of Nursing at Western University in London, ON.