BMC Medical Research Methodology

RESEARCH

Open Access

Evaluating methods to define place of residence in Canadian administrative data and the impact on observed associations with all-cause mortality in type 2 diabetes

Danielle K Nagy[1], Lauren C Bresee[2], Dean T Eurich[3] and Scot H Simpson[4*]

Abstract

Purpose An individual's location of residence may impact health, however, health services and outcomes research generally use a single point in time to define where an individual resides. While this estimate of residence becomes inaccurate when the study subject moves, the impact on observed associations is not known. This study quantifies the impact of different methods to define residence (rural, urban, metropolitan) on the association with all-cause mortality.

Methods A diabetes cohort of new metformin users was identified from administrative data in Alberta, Canada between 2008 and 2019. An individual's residence (rural/urban/metropolitan) was defined from postal codes using 4 different methods: residence defined at 1-year before first metformin (this served as the reference model), comparison 1- stable residence for 3 years before first metformin, comparison 2– residence as time-varying (during the outcome observation window), and comparison 3 - nested case control (residence closest to the index date after identifying cases and controls). Multivariable Cox proportional hazard and logistic regression models were constructed to examine the association between residence definitions and all-cause mortality.

Results We identified 157,146 new metformin users (mean age of 55 years and 57% male) and 8,444 (5%) deaths occurred during the mean follow up of 4.7 (SD 2.3) years. There were few instances of moving after first metformin; 2.6% of individuals moved to a smaller centre (metropolitan to urban or rural, or urban to rural) and 3.1% moved to a larger centre (rural to urban or metropolitan, or urban to metropolitan). The association between rural residence and all-cause mortality was consistent (aHR:1.18; 95%CI:1.12–1.24), regardless of the method used to define residence.

Conclusions The method used to define residence in a population of adults newly treated with metformin for type 2 diabetes has minimal impact on measures of all-cause mortality, possibly due to infrequent migration. The observed

Prior presentations: This study was presented at the International Society for Pharmacoepidemiology Annual Conference as a poster presentation. Conference dates August 24–28, 2024.

*Correspondence:
Scot H Simpson
scot@ualberta.ca

Full list of author information is available at the end of the article

association between residence and mortality is compelling but requires further investigation and more robust analysis.

Plain language summary

There is growing evidence describing the impact of where an individual lives on their health. However, most of these studies identify place of residence at a single point in time and do not consider when a person moves. This could result in misclassification, which could over- or underestimate the influence of residence on health outcomes. In this study, residence was defined with 4 different methods: at 1-year before starting metformin for type 2 diabetes; stable residence for 3-years before starting metformin; accounting for changes in residence after being newly treated with metformin for type 2 diabetes; and near the end of the study at death or the end of follow up. The results of this study describe that individuals rarely move after treatment initiation with metformin for type 2 diabetes and that living in a rural area has a higher risk of death from any cause, further investigation into the latter is required.

Keypoints

- Individuals rarely move after treatment initiation with metformin for type 2 diabetes.
- In population-based cohort studies of adults with type 2 diabetes, classifying place of residence at baseline, 1-year prior to the index date is reasonable.
- The increased mortality of rural residents requires further investigation.

Keywords Rural-urban continuum, Migration, Type 2 diabetes, Time-fixed, Time-varying, Nested case-control

Introduction

Across various disease states, there is mounting evidence of the impact of an individual's residence on their health [1–22]. In Canada, it is reported that rural dwellers with diabetes are up to 50% more likely to experience a diabetes-related avoidable hospitalization than urban dwellers [1, 5]. Similar findings of increased diabetes-related morbidity and mortality are reported in rural areas of the United States as well [6, 7]. Rural-urban disparities have also been described with respect to barriers to participating in physical activity, treatment outcomes for obstructive sleep apnea, access to pediatric trauma centres, the epidemiology of different cancers, the prevalence of osteoarthritis-related comorbidities, and gradients in perinatal health.[11–14,16–20,22] Even mainstream media has reported place of residence as an important social determinant of health, suggesting an individual's "ZIP code is as important as their genetic code." [23].

From an epidemiologic perspective, there are several ways to classify where an individual lives. In the United States, some measures of residence are based off of census-derived data and reported at the state or county level [10, 24, 25]. The United States also has rural-urban commuting area codes to classify census tracts as metropolitan, micropolitan, small town, and rural commuting areas based on population density, urbanization, and daily commuting [26]. In Canada, census-derived measures of residence are also used, however, provincial jurisdictions, responsible for the delivery of healthcare, developed their own definitions of residence for healthcare reporting and surveillance [27–29]. Provincial jurisdictions base their residence boundaries off of population density, local industry, travel patterns, infrastructure, and more, and are reported at the individual-level in administrative datasets used for research [28, 30, 31].

There have been several studies comparing the geographic accuracy of different residence definitions and the impact on the prevalence of different diseases and disease-related morbidity and mortality [10, 17, 19, 20, 25, 27, 32–39]. However, most identify residence at cohort entry or cross-sectionally at a single point in time and fail to consider the impact of migration which may lead to error [40, 41]. Another limitation of health mobility research is the assumption that time has a "constant and uniform effect" which does not consider critical periods around *when* a move occurs, but rather only *if* a move occurs [41]. The purpose of this study was therefore to use methods to define residence that would address these limitations and determine if these different methods would affect the association with all-cause mortality. A population of individuals with type 2 diabetes was selected for this study based on the availability of the data, duration of follow up in which migration could be observed, and previously reported disparities in this population according to residence across the rural-urban continuum [1–7]. Given the chronic and progressive nature of type 2 diabetes, it is plausible that these individuals over time would choose to move to a metropolitan or urban area with more specialized healthcare facilities and resources. The current interprovincial migration of individuals with type 2 diabetes in Alberta is unknown, yet integral to health system planning, allocation of resources, and future interventional studies.

Methods

Data source and study population

We conducted a retrospective cohort study between 1 April 2008 and 31 March 2019 using linkable administrative datasets from Alberta Health including population registry, ambulatory care, inpatient, pharmaceutical information network, practitioner claims, and vital statistics [42]. To be eligible for cohort entry, individuals were required to be an adult, eligible for healthcare benefits in Alberta, and have at least 1 dispensation for any antihyperglycemic drug therapy between the above dates, from which new metformin users were identified. A new user design was then employed which was defined as having no history of any antihyperglycemic drug therapy dispensation for at least 12 months prior to the first metformin dispensation [43]. Individuals were excluded for the following reasons: ICD-9 code 256.4 or ICD-10 code E28.2 (polycystic ovarian syndrome) at any time, ICD-10 code O24.x (gestational diabetes) 9 months before first metformin to the end of follow up, less than 1 year of follow up after first metformin dispensation, and age less than 18 years at the time of first metformin dispensation (see Fig. 1). The study index date was defined as the date of the first metformin dispensation.

Exposure definitions: location of residence

For all exposure definitions, residence was categorized as rural, urban, or metropolitan using an individual's postal code. The residence categories were based on Alberta Health Services and Alberta Health's standard geographic areas for health system planning and surveillance; residence boundaries are set according to population density, local industry, travel patterns, infrastructure, and more [28, 30, 31]. Metropolitan areas are defined by populations greater than 500,000 people (Edmonton and Calgary) and surrounding areas that are commuter communities. Urban refers to populations between 25,000 and 500,000 people. Rural areas have smaller population sizes, unique industries, and increased distances from a metropolitan or urban centre [28]. It should be noted that residence could be identified from the administrative data at each fiscal year end (March 31 of each year).

Methods that are commonly used to identify drug exposures in pharmacoepidemiology (time-fixed,

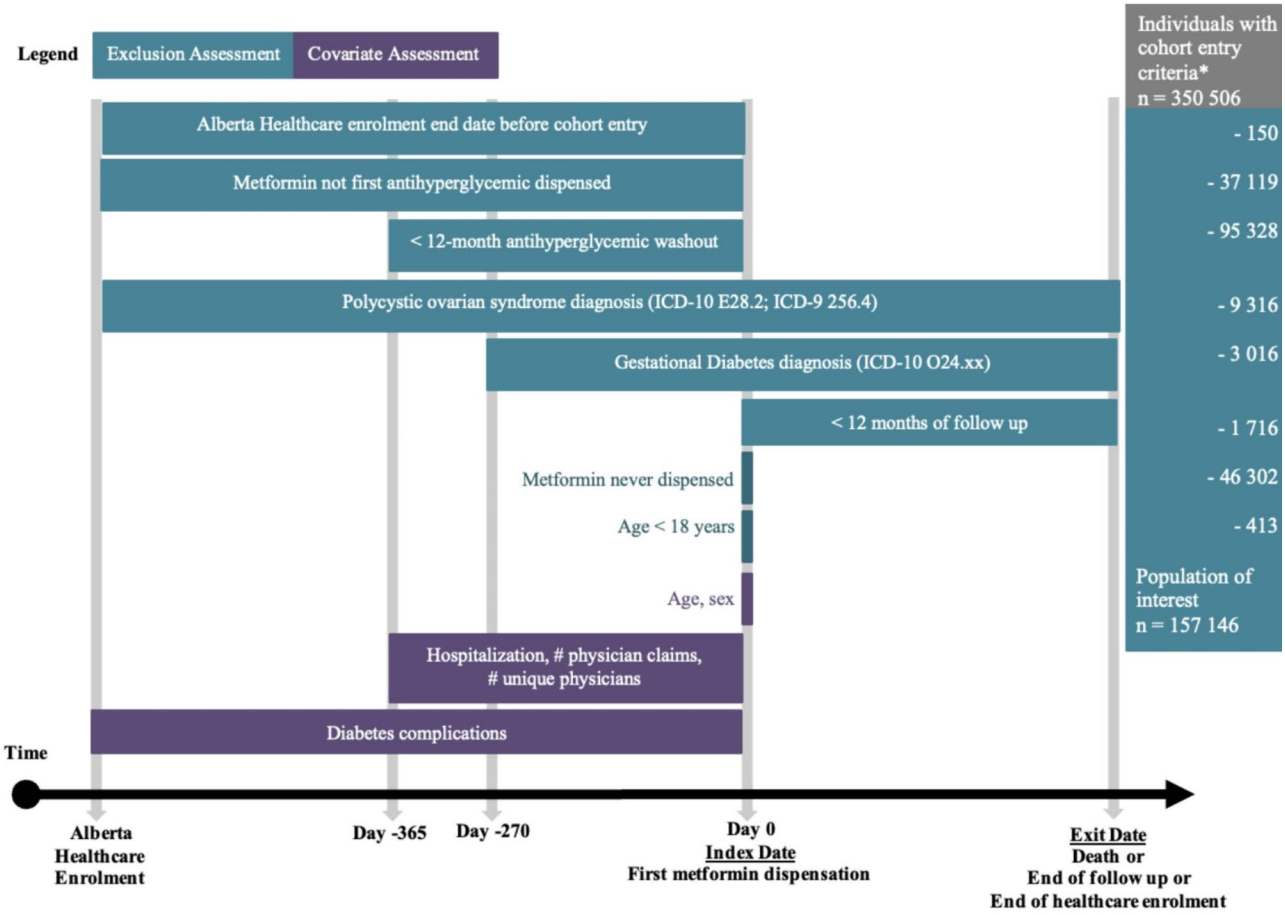

Fig. 1 Population Flow Diagram: Defining Population of Interest. Note. * Alberta residents with ≥ 1 dispensation of antihyperglycemic drug therapy between 1 April 2008 and 31 March 2019; ICD = World Health Organization International Classification of Diseases

time-varying, and nested case-control) were used to define how and when residence as an exposure was operationalized (Fig. 2) [44–48]. In the reference model, residence was defined using the closest fiscal year end at least 1-year prior to the first metformin dispensation. This time-fixed approach is commonly used in current literature but has the potential for misclassification if an individual has moved before or after this date [40, 41].

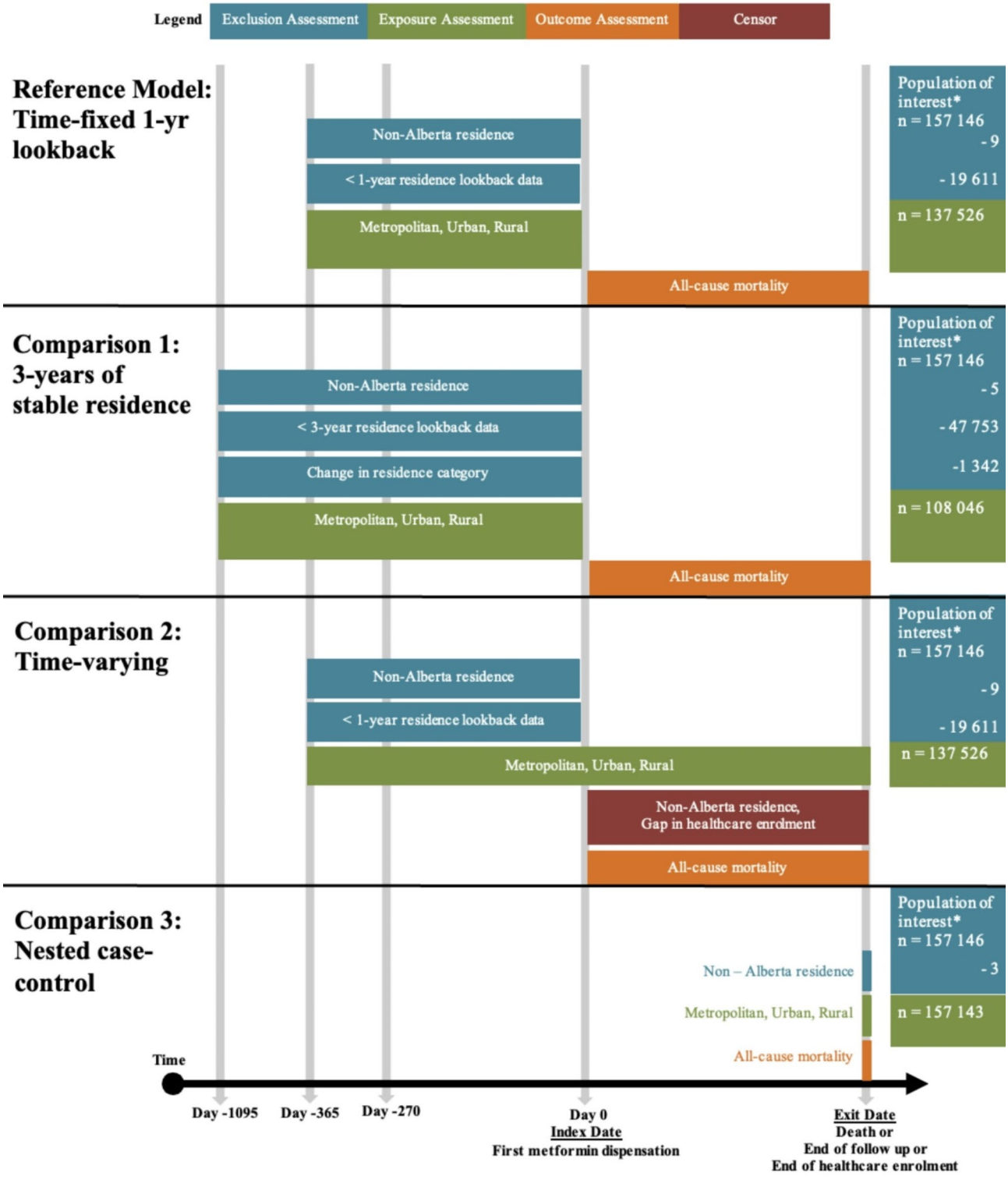

Fig. 2 Population Flow Diagram: Approaches to Defining Residence. Note. *Refer to Fig. 1 for identification of population of interest

Comparison 1 mitigates misclassification by excluding movers within 3-years before the first metformin dispensation. For example, if 1 year prior to first metformin, an individual was reported to be living in a rural area, but the year prior to that (2 years before metformin) the individual was reported to be living in an urban area, they would be included in the reference model as living in a rural area, but excluded in comparison 1 as their residence changed within the 3 years prior to metformin.

Comparison 2 mitigates misclassification of residence by allowing residence to be time-dependent. In this model, residence was initially defined at 1-year prior to first metformin then, at each subsequent fiscal year end during the observation window, the residence category was reviewed and reclassified if the individual moved (Fig. 2). This continued for each fiscal year end until the end of follow up or death occurred.

Comparison 3 considers whether residence closest to the event date impacts the outcome. Comparison 3 utilized a nested case-control approach whereby after identifying new metformin users and applying the exclusion criteria described in Fig. 1, each individual was classified as a case (having died within the observation window) or control (follow up ending before death occurred) (Fig. 2). For cases, residence was defined at the fiscal year end when death occurred. Similarly, for controls, residence was defined at the fiscal year end when follow up ceased.

Outcome: All-cause mortality

While the objective of this study was not to determine whether an association exists between residence and all-cause mortality, the outcome of all-cause mortality was chosen for being broad (not specifically attributable to a medication exposure, comorbidities, etc.) and therefore all individuals in the cohort would be at risk of the outcome. Date of death was recorded in the vital statistics dataset, provided by Service Alberta (who issues death certificates in the province of Alberta) [42].

Covariates

Each of the 4 models were adjusted for the same covariates to ensure that the only difference between the models was the exposure definition. Additionally, covariate adjustment was limited as the models were not intended to predict causality; only healthcare utilization metrics known to differ according to residence and complications known to impact diabetes-related morbidity and mortality were included. Healthcare utilization covariates were identified at baseline, in the 1-year prior to first metformin dispensation, as follows: a count of the number of physician claims (categorized as 0–6, 7–12, 13–24, and ≥ 25), whether at least 1 hospitalization occurred, and the number of unique physicians visited (categorized as 0–2, 3–4, 5–8, and ≥ 9). Cut points for categorization were chosen to be clinically meaningful and that best fit the data to address skewness. This has been further described elsewhere [2, 3]. Age at first metformin dispensation and sex assigned at birth were also controlled for in the analysis.

An adapted Diabetes Complications Severity Index (aDCSI) score was calculated for each individual as a sum of the severity of 7 diabetes complications (retinopathy, nephropathy, neuropathy, cerebrovascular, cardiovascular, peripheral vascular disease, and metabolic) [49]. Grading of severity for each complication is either 0, 1, or 2 based on the presence of ICD-9 and ICD-10 codes and does not use laboratory data [49–51]. Complications were identified at baseline, any time prior to first metformin dispensation.

Statistical analysis

For both the reference model and comparisons 1 and 2, Cox proportional hazard modelling was conducted, and individuals were followed until they experienced the outcome or follow up ceased due to the end of the study data or end of healthcare enrolment (Fig. 2). Specific to comparison 2 where residence was time-dependent, if an individual moved (defined as a change in residence classification) during the observation window, the time at risk was accounted for in each residence location [52]. A multivariable logistic regression analysis was performed for comparison 3 (the nested case control approach). All covariates previously mentioned were adjusted for in each of the 4 models; the reference group used for all models was metropolitan residence. Stata version 16.1 was used for all analyses.

Results

Baseline population characteristics

Of 350,506 Alberta residents with at least 1 dispensation of antihyperglycemic drug therapy between April 1, 2008 and March 31, 2019, 157,146 met the inclusion criteria and were identified as new metformin users (Fig. 1). Populations were further defined based on meeting different exposure definitions for residence. In the reference model and comparison 2, the populations were further reduced to 137,526 after excluding those with less than 1 year of residence data. Comparison 1 and 3 resulted in populations of 108,046 and 157,143, respectively (Fig. 2).

Residence was distributed as 26% rural, 10% urban, and 64% metropolitan for all exposure definitions. In the population defined using the time-fixed, 1-year lookback approach, the mean age was 55 years, 57% were male, and the mean follow up time was 4.7 years (Table 1). The largest differences in baseline characteristics were between metropolitan and rural populations in the number of physician claims in the year prior to first metformin and the number of unique physicians visited in the same time

Table 1 Baseline demographics by place of residence at 1-year lookback from first Metformin dispensation

	Metropolitan (*n* = 87,778)	Urban (*n* = 13,945)	Rural (*n* = 35,803)	Maximum Absolute Standardized Difference
Age (years), mean (SD)	55 (13.6)	54 (13.6)	56 (13.8)	0.19 (U-R)
Male, n (%)	49,654 (56.6)	8,135 (58.3)	20,456 (57.1)	0.04 (M-U)
Number of physician claims, n (%)				0.20 (M-R)
0–6	23,714 (27.0)	4,769 (34.2)	12,836 (35.9)	
7–12	22,482 (25.6)	3,590 (25.7)	8,907 (24.9)	
13–24	22,528 (25.7)	3,184 (22.8)	8,001 (22.3)	
≥ 25	19,054 (21.7)	2,402 (17.2)	6,059 (16.9)	
Number of unique physicians visited, n (%)				0.34 (M-R)
0–2	16,341 (18.6)	3,682 (26.4)	11,151 (31.1)	
3–4	22,548 (25.7)	3,894 (27.9)	9,943 (27.8)	
5–8	28,904 (32.9)	4,012 (28.8)	9,483 (26.5)	
≥ 9	19,985 (22.8)	2,357 (16.9)	5,226 (14.6)	
Hospitalization, n (%)	9,401 (10.7)	1,714 (12.3)	5,267 (14.7)	0.12 (M-R)
aDCSI Score, mean (SD)	0.8 (1.3)	0.8 (1.3)	0.9 (1.4)	0.09 (M-R)

Note. SD = standard deviation, M-U = metropolitan-urban comparator, U-R = Urban-rural comparator, M-R = metropolitan-rural comparator, aDCSI = adapted Diabetes Complications Severity Index Score as described in Chang HY, Weiner JP, Richards TM, et al. Validating the adapted Diabetes Complications Severity Index in claims data. *Am J Manag Care* 2012;18(11):721-6

Table 2 Effect of residence on all-cause mortality using different exposure definitions

		Deaths, n (%) [Incidence rate per 1000 person years]			Measure of Association* (95% Confidence Interval)	
		Rural	Urban	Metropolitan	Rural vs. Metropolitan	Urban vs. Metropolitan
Residence Exposure Definition	Reference Model: Time-fixed 1-yr lookback	2,671 (7.5) [15.5]	809 (5.8) [12.3]	4,964 (5.7) [12.0]	aHR: 1.18 (1.12–1.24)	aHR: 1.14 (1.06–1.23)
	Comparison 1: 3 years of stable residence	1,861 (6.4) [14.7]	524 (4.9) [11.6]	3,275 (4.8) [11.4]	aHR: 1.17 (1.10–1.24)	aHR: 1.08 (0.99–1.19)
	Comparison 2: Time-varying	2,623 (7.4) [15.3]	835 (6.0) [12.8]	4,986 (5.7) [12.1]	aHR: 1.17 (1.12–1.23)	aHR: 1.17 (1.08–1.25)
	Comparison 3: Nested Case Control	3,405 (8.4)	1,083 (14.5)	6,890 (6.8)	aOR: 1.17 (1.12–1.23)	aOR: 1.12 (1.04–1.20)

Note. * Metropolitan as reference in all models and adjusted for age, sex, number of physician claims 1-year prior to first metformin, number of unique physicians visited 1-year prior to first metformin, hospitalization within 1-year prior to first metformin, and adapted Diabetes Complications Severity Index score

period, standardized difference 0.20 and 0.34, respectively (Table 1). Baseline population characteristics were similar for the other cohorts identified (Supplementary Tables 1 and Supplementary Table 2).

Moving after first Metformin using the time-varying exposure definition

When considering residence as a time-varying exposure, 7,864 out of 137,526 people (5.7%) had at least 1 record of moving (a change in residence category) after their first metformin dispensation. Of these individuals, 3,599 people (2.6%) moved to a smaller area (metropolitan to urban, metropolitan to rural, or urban to rural) and 4,265 people (3.1%) moved to a larger area (rural to urban, rural to metropolitan, or urban to metropolitan). The largest number of moves a single person experienced was 6.

All-cause mortality

Table 2 shows that the number of events by residence category was similar across all models and demonstrates a geographic gradient. Metropolitan had the fewest deaths in all models with incidence rates ranging from 11.4 to 12.1 per 1000 person years, followed by urban (11.6 to 12.8 per 1000 person years), and then rural (14.7 to 15.5 per 1000 person years).

The effect of each residence exposure definition on all-cause mortality was similar across models and rural residence was associated with increased mortality in all models (Table 2). The adjusted hazard ratio for all-cause mortality among rural residents was 1.18; 95% CI:1.12–1.24 in the reference model and 1.17; 95% CI:1.10–1.24 in comparison 1 when misclassification was mitigated by excluding movers in the 3 years prior to first metformin. When allowing residence to be time-dependent

(comparison 2) the adjusted hazard ratio for all-cause mortality was 1.17; 95% CI:1.12–1.23 for rural residents. Considering where an individual resided at the time of death or end of follow up (comparison 3) resulted in an adjusted odds ratio of 1.17; 95% CI:1.12–1.23 for rural residents.

Discussion

This study describes the migration of adult new metformin users across the rural-urban continuum in Alberta, Canada and quantifies the impact of different methods to define residence on measures of all-cause mortality. The results of this study demonstrate that individuals rarely move between residence categories (rural, urban, metropolitan) after becoming new metformin users and of those who do move, the proportion of those moving to a larger or smaller area are similar. The primary implication of this study is that the method used to define residence, for use in statistical modelling, does not substantially impact the measure of association. In situations with minimal migration, using residence as a time-varying exposure or covariate appears to be unnecessary. Rather, classifying residence at baseline, at 1-year before the index date in populations of adults newly treated with metformin for type 2 diabetes, is acceptable.

There is little else described in the literature on migration patterns of individuals with type 2 diabetes in Canada or the United States. The limited migration uncovered in this study is however, supported by literature in other disease states and Canadian provinces. In a population of Albertans with Parkinson's disease or multiple sclerosis, migration within the province following disease onset was minimal and moving within the same municipality accounted for most (80%) of the observed migration [32]. Similarly, in a population of individuals with rheumatoid arthritis in the province of Quebec, most migration following diagnosis was inter-urban, not as a result of rural-to-urban migration [53]. In general, following a new diagnosis or disease onset, Canadians continue to live in the same places as before, making residence a stable metric over time.

The incidental finding of increased mortality in people residing in rural areas, compared to metropolitan Alberta is concerning. While the statistical models presented in this study were not intended to describe a causal relationship, this is an important signal which requires further investigation. We have previously described differential management strategies undertaken for type 2 diabetes in rural Alberta, which we hypothesize may impact morbidity and mortality [2–4]. Similarly, others have described a substantially higher burden of diabetes-related mortality in rural areas of the United States [6, 7]. However, despite the consideration of multiple factors such as diet, lifestyle, exposure to environmental pesticides and fertilizers, comorbidities, and healthcare infrastructure, the authors conclude that the reasons for rural-urban disparities in type 2 diabetes are still incompletely understood and further studies are required [7]. Given this evidence, more robust analyses are planned to measure the impact of residence as an exposure on mortality related to type 2 diabetes.

Strengths and limitations

This study is strengthened by the use of individual-level residence data which was linked to individual-level health data. This provides a unique perspective and deeper insight into the impact of residence on health outcomes, considering that many studies conducted elsewhere such as in the United States, may utilize census-derived residence data which is reported at the state or county-level [7, 10, 24]. That being said, we did not classify residence by city or town, to maintain patient confidentiality and anonymity, and therefore we did not account for whether an individual moved within residence categories, for example, from one rural town to another rural town. Additionally, the results of this study may not be generalizable to other populations, especially given the numerous ways to define residence. While the results of this study are geographically specific to Alberta, Canada, it does generate further testable hypotheses to other places and populations.

The approaches used in this study also only consider immediate effects, however, lagged effects of past exposure, such as place of birth or residence during formative years of development, could also impact the outcome. A longitudinal birth cohort would be required to determine these effects. Additionally, residual confounding is present in the estimates of all-cause mortality, given the limited adjustment in each model. However, these models were not intended to be causal but rather, quantify differences in the measures of association given changes to the exposure definition of residence alone. More robust models are required to determine whether there is a causal link between residence along the rural-urban continuum and mortality.

Conclusion

The method used to define residence in a population of new metformin users had minimal impact on measures of all-cause mortality, possibly owing to the infrequency with which people move within the province of Alberta. This study provides reasonable justification for classifying place of residence in population-based cohort studies at baseline and at 1-year before the index date, compared to other methods, provided that migration within the cohort is minimal. Incidental findings of increased mortality in rural areas are being further explored to discern whether there is a causal relationship.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12874-025-02531-3.

Supplementary Material 1

Acknowledgements
This study is based in part on data provided by Alberta Health. The interpretation and conclusions contained herein are those of the researchers and do not necessarily represent the views of the Government of Alberta. Neither the Government of Alberta nor Alberta Health express any opinion in relation to this study.

Author contributions
All authors were responsible for the concept and design of the study. S.H.S. and D.T.E. were responsible for data acquisition. D.K.N. performed the analysis, and all authors interpreted the results. D.K.N. prepared an initial draft of the manuscript. All authors helped revise the manuscript as needed.

Funding
S.H.S. is supported as the Chair in Patient Health Management, jointly held by the Faculty of Pharmacy and Pharmaceutical Sciences and Faculty of Medicine and Dentistry, University of Alberta.
This study was not supported by a grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability
The administrative health data used to support the findings of this study were obtained from Alberta Health under a data disclosure agreement and are not publicly available. Requests for administrative health data can be submitted to Alberta Health (https://www.alberta.ca/health-research).

Declarations

Ethics approval and consent to participate
The University of Alberta Research Ethics Board approved the conduct of this study (Pro00066037). Individual level consent was waved as this study used de-identified administrative health data and disclosure for research purposes followed regulations of the Alberta Health Information Act (https://open.alberta.ca/publications/h05).

Consent for publication
Not Applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, 2-35, Medical Sciences Building, 8613– 114 St, Edmonton, AB T6G1C9, Canada
[2]Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, 3280 Hospital Drive NW, Calgary, AB T2N4Z6, Canada
[3]School of Public Health, University of Alberta, 2-040F Li Ka Shing Centre For Research, 11203– 87 Ave NW, Edmonton, AB T6G2H5, Canada
[4]Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, 2-020C, Katz Group Centre for Research, 11315– 87 Ave NW, Edmonton, AB T6G2H5, Canada

Received: 4 November 2024 / Accepted: 10 March 2025
Published online: 18 March 2025

References

1. Gupta N, Crouse DL, Miah P, Takaro T. The role of neighbourhood environments in hospitalization risk for diabetes and related conditions: A population-based cohort analysis by remoteness and deprivation indices. Health Rep Dec. 2022;21(12):3–13. https://doi.org/10.25318/82-003-x202201200001.
2. Nagy DK, Bresee LC, Eurich DT, Simpson SH. Are guideline concordant processes of care consistent across the Rural-Urban continuum?? A retrospective cohort study of adults newly treated for type 2 diabetes. Can J Diabetes. 2024;48(5):322–e3295. https://doi.org/10.1016/j.jcjd.2024.03.007.
3. Nagy DK, Bresee LC, Eurich DT, Simpson SH. Rural residence is associated with a delayed trend away from sulfonylurea use for treatment intensification of type 2 diabetes. Diabetes Care Mar. 2023;01(3):613–9. https://doi.org/10.2337/dc22-1223.
4. Nagy DK, Bresee LC, Eurich DT, Simpson SH. Rurality is associated with lower likelihood of dipeptidyl peptidase 4 inhibitor use for treatment intensification. Explor Res Clin Soc Pharm Mar. 2024;13:100429. https://doi.org/10.1016/j.rcsop.2024.100429.
5. Booth GL, Hux JE, Fang J, Chan BT. Time trends and geographic disparities in acute complications of diabetes in Ontario, Canada. Diabetes Care May. 2005;28(5):1045–50. https://doi.org/10.2337/diacare.28.5.1045.
6. Dugani SB, Wood-Wentz CM, Mielke MM, Bailey KR, Vella A. Assessment of disparities in diabetes mortality in adults in US rural vs nonrural counties, 1999–2018. JAMA Netw Open Sep. 2022;01(9):e2232318. https://doi.org/10.1001/jamanetworkopen.2022.32318.
7. Dugani SB, Mielke MM, Vella A. Burden and management of type 2 diabetes in rural united States. Diabetes Metab Res Rev Jul. 2021;37(5):e3410. https://doi.org/10.1002/dmrr.3410.
8. Hill-Briggs F, Adler NE, Berkowitz SA, et al. Social determinants of health and diabetes: A scientific review. Diabetes Care. 2021;44(January 2021):258–79. https://doi.org/10.2337/dci20-0053.
9. Mather H, Ornstein KA, McDonough C, Liu B, Salinger M, Ritchie CS. County-Level social vulnerability, metropolitan status, and availability of home health services. JAMA Netw Open Oct. 2023;02(10):e2337508. https://doi.org/10.1001/jamanetworkopen.2023.37508.
10. Werner AK, Strosnider H, Kassinger C, Shin M, Workgroup S-CDP. Lessons learned from the environmental public health tracking Sub-County data pilot project. J Public Health Manag Pract. 2018;24(5):E20–7. https://doi.org/10.1097/PHH.0000000000000686.
11. Corrigan J, Ogah I, Ip-Buting A, et al. An evaluation of rural-urban disparities in treatment outcomes for obstructive sleep apnoea: study protocol for a prospective cohort study. ERJ Open Res Oct. 2020;6(4). https://doi.org/10.1183/23120541.00141-2020.
12. Amram O, Schuurman N, Pike I, Friger M, Yanchar NL. Assessing access to paediatric trauma centres in Canada, and the impact of the golden hour on length of stay at the hospital: an observational study. BMJ Open Jan. 2016;08(1):e010274. https://doi.org/10.1136/bmjopen-2015-010274.
13. Pelletier CA, White N, Duchesne A, Sluggett L. Barriers to physical activity for adults in rural and urban Canada: A cross-sectional comparison. SSM Popul Health Dec. 2021;16:100964. https://doi.org/10.1016/j.ssmph.2021.100964.
14. Pelletier C, White N, Duchesne A, Sluggett L. Rural-urban differences in individual and environmental correlates of physical activity in Canadian adults. Prev Med Rep Dec. 2022;30:102061. https://doi.org/10.1016/j.pmedr.2022.102061.
15. St John PD, Menec V, Tate R, Newall NE, Cloutier D, O'Connell M. Depressive symptoms in adults in rural and urban regions of Canada: a cross-sectional analysis of the Canadian longitudinal study on aging. BMJ Open Dec. 2021;03(12):e048090. https://doi.org/10.1136/bmjopen-2020-048090.
16. Roy SF, Ghazawi FM, Le M, et al. Epidemiology of adult and pediatric Burkitt lymphoma in Canada: sequelae of the HIV epidemic. Curr Oncol Apr. 2020;27(2):83–9. https://doi.org/10.3747/co.27.5775.
17. Cattelan L, Ghazawi FM, Le M, et al. Geographic and socioeconomic disparity of gastric cancer patients in Canada. Curr Oncol May. 2021;28(3):2052–64. https://doi.org/10.3390/curroncol28030190.
18. Fakanye O, Singh H, Desautels D, Torabi M. Geographical variation and factors associated with gastric cancer in Manitoba. PLoS ONE. 2021;16(7):e0253650. https://doi.org/10.1371/journal.pone.0253650.
19. Liu X, Shahid R, Patel AB, et al. Geospatial patterns of comorbidity prevalence among people with osteoarthritis in Alberta Canada. BMC Public Health Oct. 2020;15(1):1551. https://doi.org/10.1186/s12889-020-09599-0.
20. Liu X, Seidel JE, McDonald T, et al. Rural-Urban differences in Non-Local primary care utilization among people with osteoarthritis: the role of Area-Level factors. Int J Environ Res Public Health May. 2022;24(11). https://doi.org/10.3390/ijerph19116392.
21. Liu X, Seidel JE, McDonald T, et al. Rural-Urban disparities in realized Spatial access to general practitioners, orthopedic surgeons, and physiotherapists

among people with osteoarthritis in Alberta, Canada. Int J Environ Res Public Health Jun. 2022;23(13). https://doi.org/10.3390/ijerph19137706.
22. Ospina M, Osornio-Vargas Á, Nielsen CC, et al. Socioeconomic gradients of adverse birth outcomes and related maternal factors in rural and urban Alberta, Canada: a concentration index approach. BMJ Open Feb. 2020;02(1):e033296. https://doi.org/10.1136/bmjopen-2019-033296.
23. Foden-Vencil K, Your ZIP, Code Might Be As Important To Health As Your Genetic Code. NPR. Health Reporting in the States. Accessed July 3, 2024, https://www.npr.org/sections/health-shots/2017/08/04/539757759/your-zip-code-can-be-as-important-to-health-as-your-genetic-code
24. NORC. The Mapping Medicare Disparities Tool. Technical Documentation Centers for Medicare & Medicaid Services. Office of Minority Health. Accessed July 2. 2024, https://www.cms.gov/priorities/health-equity/minority-health/research-data/mapping-medicare-disparities-tool-mmd
25. Werner AK, Strosnider HM. Developing a surveillance system of sub-county data: finding suitable population thresholds for geographic aggregations. Spat Spatiotemporal Epidemiol Jun. 2020;33:100339. https://doi.org/10.1016/j.sste.2020.100339.
26. USDA Economic Research Service. US Department of Agriculture. Rural-Urban Commuting Area Codes. Updated 1/6/2025. Available from [https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes] Accessed February 23, 2025.
27. Pinault L, Khan S, Tjepkema M. Accuracy of matching residential postal codes to census geography. Health Rep Jun. 2020;17(3):3–13. https://doi.org/10.25318/82-003-x202000300001-eng.
28. Alberta Health Services and Alberta Health. Official Standard Geographic Areas. Alberta Government. 2018. Accessed August 5, 2024. https://open.alberta.ca/publications/official-standard-geographic-areas
29. Statistics Canada. Postal Code OM Conversion File Plus (PCCF+). Government of Canada. Accessed July 3. 2024. https://www150.statcan.gc.ca/n1/en/catalogue/82F0086X
30. Government of Alberta. Overview of Administrative Health Datasets. Alberta Health Analytics and Performance Reporting Branch. Accessed April 12. 2023, https://open.alberta.ca/dataset/657ed26d-eb2c-4432-b9cb-0ca2158f165d/resource/38f47433-b33d-4d1e-b959-df312e9d9855/download/research-health-datasets.pdf
31. Alberta Health. *Postal Code Translator File (PCTF)*. September 2021.
32. Yiannakoulias N, Schopflocher DR, Warren SA, Svenson LW. Parkinson's disease, multiple sclerosis and changes of residence in Alberta. Can J Neurol Sci Aug. 2007;34(3):343–8. https://doi.org/10.1017/s0317167100006806.
33. Bow CJ, Waters NM, Faris PD, et al. Accuracy of City postal code coordinates as a proxy for location of residence. Int J Health Geogr Mar. 2004;18(1):5. https://doi.org/10.1186/1476-072X-3-5.
34. Fuller D, Shareck M. Canada post community mailboxes: implications for health research. Can J Public Health Sep. 2014;17(6):e453–5. https://doi.org/10.17269/cjph.105.4721.
35. Jaworsky D, Loutfy M, Lu M, et al. Influence of the definition of rurality on geographic differences in HIV outcomes in British Columbia: a retrospective cohort analysis. CMAJ Open Oct-Dec. 2020;8(4):E643–50. https://doi.org/10.9778/cmajo.20200066.
36. Terashima M, Kephart G. Misclassification errors from postal code-based geocoding to assign census geography in Nova Scotia, Canada. Can J Public Health Dec. 2016;27(4–5):e424–30. https://doi.org/10.17269/cjph.107.5459.
37. Southern DA, Ghali WA, Faris PD, et al. Misclassification of income quintiles derived from area-based measures. A comparison of enumeration area and forward sortation area. Can J Public Health Nov-Dec. 2002;93(6):465–9.
38. Healy MA, Gilliland JA. Quantifying the magnitude of environmental exposure misclassification when using imprecise address proxies in public health research. Spat Spatiotemporal Epidemiol Apr. 2012;3(1):55–67. https://doi.org/10.1016/j.sste.2012.02.006.
39. Glazier RH, Creatore MI, Agha MM, Steele LS, Group ICTTTW. Socioeconomic misclassification in Ontario's health care registry. Can J Public Health. 2003;94(2):140–3. https://doi.org/10.1007/BF03404588.
40. Bentham G. Migration and morbidity: implications for geographical studies of disease. Soc Sci Med. 1988;26(1):49–54. https://doi.org/10.1016/0277-9536(88)90044-5.
41. Morris T, Manley D, Sabel CE. Residential mobility: towards progress in mobility health research. Prog Hum Geogr Feb. 2018;42(1):112–33. https://doi.org/10.1177/0309132516649454.
42. Government of Alberta. Health research data access. Accessed August 5. 2024, https://www.alberta.ca/health-research.aspx
43. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol Nov. 2003;1(9):915–20. https://doi.org/10.1093/aje/kwg231.
44. Rassen JA, Bartels DB, Schneeweiss S, Patrick AR, Murk W. Measuring prevalence and incidence of chronic conditions in claims and electronic health record databases. Clin Epidemiol. 2019;11:1–15. https://doi.org/10.2147/CLEP.S181242.
45. Eskin M, Eurich DT, Simpson SH. Impact of drug exposure definitions on observed associations in pharmacoepidemiology research. J Popul Ther Clin Pharmacol Mar. 2018;20(1):e39–52. https://doi.org/10.22374/1710-6222.25.1.4.
46. Patorno E, Patrick AR, Garry EM, et al. Observational studies of the association between glucose-lowering medications and cardiovascular outcomes: addressing methodological limitations. Diabetologia Nov. 2014;57(11):2237–50. https://doi.org/10.1007/s00125-014-3364-z.
47. Schneeweiss S, Patorno E. Conducting Real-world evidence studies on the clinical outcomes of diabetes treatments. Endocr Rev Sep. 2021;28(5):658–90. https://doi.org/10.1210/endrev/bnab007.
48. Connolly JG, Schneeweiss S, Glynn RJ, Gagne JJ. Quantifying bias reduction with fixed-duration versus all-available covariate assessment periods. Pharmacoepidemiol Drug Saf May. 2019;28(5):665–70. https://doi.org/10.1002/pds.4729.
49. Chang HY, Weiner JP, Richards TM, Bleich SN, Segal JB. Validating the adapted diabetes complications severity index in claims data. Am J Manag Care Nov. 2012;18(11):721–6.
50. Glasheen WP, Renda A, Dong Y. Diabetes complications severity index (DCSI)-Update and ICD-10 translation. J Diabetes Complications Jun. 2017;31(6):1007–13. https://doi.org/10.1016/j.jdiacomp.2017.02.018.
51. Young BA, Lin E, Von Korff M, et al. Diabetes complications severity index and risk of mortality, hospitalization, and healthcare utilization. Am J Manag Care Jan. 2008;14(1):15–23.
52. StataCorp. Stata 18 survival analysis reference manual. Stata; 2023.
53. Labrecque JA, Kyle RP, Joseph L, Bernatsky S. Health-selective migration among patients with rheumatoid arthritis in Québec: a cohort study using administrative data. Rheumatol Int Sep. 2016;36(9):1275–9. https://doi.org/10.1007/s00296-016-3540-1.

Publisher's note