In Canada, health inequities have been well documented in rural, remote and northern areas. As a result, improving rural health is often a priority area for research and service delivery. For example, rural, remote and northern areas in Canada have been found to have higher rates of all-cause mortality, chronic kidney disease, obesity, cerebrovascular disease and maternal morbidity.\(^1\)\(^-\)\(^4\) In addition to higher incidence and prevalence of chronic diseases, disease-specific outcomes and use of evidence-based therapies are poorer in rural areas.\(^5\)\(^-\)\(^7\)

Infection with HIV is a chronic disease with many reported geographic discrepancies in care in high-income countries. With appropriate care and medications, the life expectancy of those living with HIV can approach that of the general population.\(^8\)\(^-\)\(^14\) Despite advances in HIV care, people in rural areas experience more advanced disease at diagnosis, delayed linkage to care, more rapid disease progression and increased mortality.\(^9\)\(^-\)\(^14\)

The relation between rurality and health outcomes is complex and is driven by intersecting determinants, including poverty, education, industries that extract natural resources, colonialism, health human resources and travel distances.\(^15\)\(^-\)\(^21\) Further complicating health analyses on rurality are heterogeneity among rural areas, varying definitions of rurality and authors’ failure to justify selected definitions.\(^22\)\(^,\)\(^23\) In Canada, common definitions rely on population census categorizations, postal codes or rurality indices.\(^24\)\(^-\)\(^26\)

**Abstract**

**Background:** Improving rural health is often identified as a priority area for research and policy in Canada. We examined how findings on HIV outcomes (virologic suppression) can vary depending on the definition of rurality used.

**Methods:** We performed retrospective cohort analyses using the Comparative Outcomes and Service Utilization Trends study population-based cohort of adults (age ≥ 19 yr) living with HIV in British Columbia between Apr. 1, 2012, and Mar. 31, 2013. We performed univariate logistic regression analyses using the following geographic variables to predict HIV virologic suppression: rurality defined by forward sortation area, by Statistical Area Classification and by health authority. We mapped suppression using geographic information systems.

**Results:** Virologic suppression was observed in 5605 (65.2%) of 8598 participants. In univariate analysis, rurality defined by Statistical Area Classification (odds ratio [OR] 0.73, 95% confidence interval [CI] 0.65–0.82), but not by forward sortation area, was associated with lower odds of suppression. When we examined suppression by health authority, Northern Health had the lowest odds of suppression (OR 0.46, 95% CI 0.36–0.58 compared to Vancouver Coastal Health). Geographic information systems mapping showed poorer suppression in northern areas.

**Interpretation:** Health outcome findings can vary depending on the definition of the geographic variable. When including geographic variables, researchers should carefully consider variable definitions and whether other classification systems, such as north–south, are more appropriate than rurality for their analysis.
Using HIV infection in British Columbia as an example, we aimed in this exploratory work to highlight challenges with existing definitions of rurality by presenting parallel analyses using multiple rurality definitions when examining for geographic patterns of HIV virologic suppression.

Methods

Design and setting
The Comparative Outcomes and Service Utilization Trends (COAST) study is a population-based cohort of adults (≥19 yr) including all those known to be living with HIV in BC. This cohort includes deidentified health data from a linkage between several data sets hosted by Population Data BC (population-level health administrative data) and the BC Centre for Excellence in HIV/AIDS. The combined data set included geographic data on location of residence, age and sex from the Consolidation File of Population Data BC, and data on HIV laboratory testing, ethnicity and injection drug use from the Drug Treatment Program database of the BC Centre for Excellence in HIV/AIDS. We used the following Population Data BC data holdings in the HIV case-finding algorithm: Discharge Abstract Database, Pharmacet, Vital Statistics Deaths and Medical Services Plan. We performed individual-level linkages using manual, deterministic and probabilistic matching facilitated by the personal health number, a unique lifetime identifier assigned to BC residents.

Study population
This study included longitudinal data on all people living with HIV in the COAST cohort who were alive as of Apr. 1, 2012. We identified eligible people through the Drug Treatment Program registry or a validated HIV case-finding algorithm that selected people with at least 1 inpatient or 3 or more outpatient International Classification of Diseases, 9th Revision or International Classification of Diseases and Related Health Problems, 10th Revision codes associated with HIV in the Population Data BC administrative data set. This case-finding algorithm was found to have a sensitivity of 88% in previous studies using BC health administrative data. Data on HIV viral load were included for the study population between Apr. 1, 2012, and Mar. 31, 2013.

Rurality and geography definitions
We assessed 3 definitions of rurality and geography: forward sortation area, Statistical Area Classification and health authority. We assessed Statistical Area Classification as both an ordinal and a binary variable, resulting in 4 analyses.

The first 3 characters in a Canadian postal code constitute a forward sortation area. The second digit of the forward sortation area can be used in an urban–rural classification system. If the second digit is 0, it is considered rural; otherwise, it is considered urban. This binary classification system was used as the geographic predictor in the first analysis.

The Statistical Area Classification is a system developed by Statistics Canada that is frequently used to allocate communities to 1 of 7 categories by size and metropolitan influence. The 7 categories are as follows: 1) census metropolitan areas have a population of 100 000 or more, with 50 000 or more people living in a core area; 2) census agglomerations have a core population of 10 000 or more, and when the core area has more than 50 000 people, the area is divided into census tracts; 3) census agglomerations without tracts are those with 50 000 people or less; 4) strong metropolitan influenced zones are areas where 5.0%–29.9% of the labour force commutes to a census metropolitan area or census agglomeration; 5) moderate metropolitan influenced zones are areas where 0.1%–4.9% of the labour force commutes to a census metropolitan area or census agglomeration; and 6) weak metropolitan influenced zones are areas where none of the labour force commutes to a census metropolitan area or census agglomeration. In the second analysis, the 7 Statistical Area Classification categories were included as ordinal categories of the geographic predictor variable.

In the third analysis, we defined census metropolitan areas as urban and grouped the other 6 categories into 1 category to represent rural areas and small towns. We selected this division because most HIV specialty clinics and HIV-specific community organizations are located in census metropolitan areas, and many isolated communities in northern BC are considered census agglomerations but have limited access to tertiary health centres.

British Columbia is divided into 5 health authorities (Vancouver Coastal Health, Island Health, Interior Health, Fraser Health and Northern Health) that are responsible for the administration and delivery of health care services in the respective regions. The First Nations Health Authority provides services for Indigenous peoples across BC irrespective of location. Those served by the First Nations Health Authority all have a geographic location within 1 of the other health authorities and also receive services from their geographic health authority. In the fourth analysis, we used health authority as the geographic predictor.

Primary outcome
We selected HIV virologic suppression as the primary outcome because it is easily measured and represents the treatment goal for HIV management. People were categorized into a binary outcome of maintaining or not maintaining virologic suppression during the study period (Apr. 1, 2012–Mar. 31, 2013). We considered virologic suppression to be maintained if all viral load measurements during the study period were less than 200 copies/mL, or if a person began antiretroviral therapy during the study period and at least 1 viral load measurement after the antiretroviral therapy start date and all subsequent viral load measurements were less than 200 copies/mL. People who did not have any viral load measurements during the study period were categorized as not maintaining virologic suppression.
Covariate selection
Covariate selection for rural health analyses can be challenging, as rurality is a construct with no uniformly accepted definition and multiple aspects. Consequently, there are many different mechanisms by which rurality may be associated with poor health, and analyses that treat additional variables as potential confounders of this relation may be adjusting for variables that mediate, rather than confound, the relation between rurality and health outcomes (see Figure 1 for conceptual framework). We did not hypothesize a causal relation between rurality and HIV outcomes; rather, the goal was to describe geographic patterns in HIV outcomes and examine how the definition of rurality changes the odds ratio (OR) associated with the outcome. Accordingly, a confounder adjusted model is not appropriate, and only unadjusted ORs are presented.

Geographic information system analysis
We performed mapping using QGIS 3.10.1 (QGIS Geographic Information System). We accessed open-access shapefiles for forward sortation area, health authorities and health service delivery areas for BC from the Statistics Canada 2016 census boundary files and the BC Health Service Delivery Area geographic data set.40,41 We mapped HIV virologic suppression data by health service delivery area, as smaller geographic units resulted in cell sizes less than 5 and could not be presented owing to confidentiality concerns. We used colour gradients to illustrate HIV virologic suppression data by quartile.

Statistical analysis
We reported descriptive statistics by outcome (virologic suppression or no virologic suppression) for each geographic

![Figure 1: Conceptual framework of intersecting determinants of rural health.](image-url)
category and covariate. All $p$ values were reported as 2-sided. We reported categoric variables as frequencies and proportions, and compared them between geographic categories using $\chi^2$ testing. We reported continuous variables as medians and interquartile ranges, and compared them using the Kruskal–Wallis test.

We developed logistic regression models for each geographic categorization type (forward sortation area, Statistical Area Classification type as an ordinal variable and as a binary variable, and health authority). We verified assumptions for logistic regression including model fit, no overspecification of model and absence of influential outliers. Individual unadjusted ORs are presented with 95% confidence intervals (CIs). We conducted statistical analyses using SAS v. 9.4 (SAS Institute).

Ethics approval

The COAST study received approval from the University of Toronto Health Sciences Research Ethics Board, the harmonized University of British Columbia/Providence Health Care Research Ethics Board and the Simon Fraser University Research Ethics Board. The present study complies with the BC Freedom of Information and Protection of Privacy Act and did not require informed consent as it was conducted retrospectively using anonymized data.

Results

A total of 8598 people living with HIV were included in these analyses (Figure 2), of whom 5605 (65.2%) maintained virologic suppression throughout the study period (Table 1). Of the total population, 6885 (80.1%) were men. The most frequently reported ethnicity was white (2728 [31.7%]), but ethnicity was not reported for 4489 (52.2%) people. When rurality was defined by forward sortation area and Statistical Area Classification, 515 (6.0%) and 1440 (16.7%) people, respectively, resided in rural areas. Most people (4623 [53.8%]) resided within the Vancouver Coastal Health region, followed by the Fraser Health region (2038 [23.7%]).

**HIV virologic suppression by rurality definition**

Table 2 presents univariate logistic regression analyses using different definitions of rurality for the geographic variable. Rurality as defined by Statistical Area Classification (OR 0.73, 95% CI 0.65–0.82), but not when defined by forward sortation area, was associated with significantly poorer virologic suppression. Among the Statistical Area Classification types, census agglomerations were associated with the lowest virologic suppression rates, and living in more rural areas (strong, moderate, weak or no metropolitan influenced zone) did not have an association with virologic suppression. Northern Health had the lowest odds of suppression (OR 0.46, 95% CI 0.36–0.58 compared to Vancouver Coastal Health).

**Mapping rurality**

Appendix 1, Supplemental Figure S1 (available at www.cmajopen.ca/content/8/4/E643/suppl/DC1) depicts a map of BC where the province is dichotomized into rural and non-rural by forward sortation area. Six maps of BC that depict which areas are classified as rural based on varying thresholds of Statistical Area Classification type are presented in Appendix 1, Supplemental Figure S2. Appendix 1, Supplemental Figure S2A represents how rurality was dichotomized where census metropolitan areas were defined as nonrural and all other areas were defined as rural, and Supplemental Figures S2B–S2F show how rurality can be dichotomized by including more Statistical Area Classification types in the nonrural category. The different health authorities in BC are depicted in Appendix 1, Supplemental Figure S3; this figure also shows the health service delivery boundaries within each health authority and is also the base map used to present the HIV virologic suppression data (below).

**Mapping HIV virologic suppression**

Figure 3 presents a map of HIV virologic suppression rates in each health service delivery area. Rates were lower in the more northern regions of the province.

**Interpretation**

We found that the association between rurality and HIV virologic suppression varied depending on the definition of rurality used. Rurality defined by Statistical Area Classification, but not by forward sortation area, was associated with lower odds of suppression. We found a stronger association and improved model fit (c-statistic) when using health authority as the geographic variable; with this
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. (%) of people*</th>
<th>Not suppressed</th>
<th>Suppressed</th>
<th>p value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>878 (51.2)</td>
<td>835 (48.7)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td>2115 (30.7)</td>
<td>4770 (69.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Age, median (IQR), yr</strong></td>
<td></td>
<td>46.2 (38.2–53.1)</td>
<td>48.9 (42.2–55.5)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td><strong>History of injection drug use‡</strong></td>
<td></td>
<td></td>
<td></td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>662 (18.0)</td>
<td>3026 (82.0)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td>787 (32.2)</td>
<td>1653 (67.7)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td></td>
<td>1544 (62.5)</td>
<td>926 (37.5)</td>
<td></td>
</tr>
<tr>
<td><strong>CD4 count, median (IQR), cells/mm³</strong></td>
<td></td>
<td></td>
<td></td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>End of follow-up</td>
<td>390 (210–600)</td>
<td>550 (390–720)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nadir</td>
<td>200 (70–360)</td>
<td>170 (70–270)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No. of viral load tests during study period, median (IQR)</strong></td>
<td></td>
<td>4 (2–5)</td>
<td>4 (3–4)</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>Ethnicity§</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td></td>
<td>n = 4489 (52.2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td>n = 268 (3.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td>n = 171 (1.9%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indigenous</td>
<td></td>
<td>n = 811 (9.4%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td>n = 131 (1.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td></td>
<td>n = 2728 (31.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Geographic variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rurality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward sortation area</td>
<td></td>
<td>2806 (34.7)</td>
<td>5277 (65.3)</td>
<td>0.46</td>
</tr>
<tr>
<td>Not rural</td>
<td></td>
<td>187 (36.3)</td>
<td>328 (63.7)</td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Statistical Area Classification</strong></td>
<td></td>
<td></td>
<td></td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Not rural</td>
<td>2405 (33.6)</td>
<td>4753 (66.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>588 (40.8)</td>
<td>852 (59.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Statistical Area Classification category</strong></td>
<td></td>
<td></td>
<td></td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Census metropolitan area</td>
<td>2405 (33.6)</td>
<td>4753 (66.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Census agglomeration (with tracts)</td>
<td>222 (45.1)</td>
<td>270 (54.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Census agglomeration (no tracts)</td>
<td>169 (42.5)</td>
<td>229 (57.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong metropolitan influenced zone</td>
<td>24 (29.3)</td>
<td>58 (70.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate metropolitan influenced zone</td>
<td>77 (34.1)</td>
<td>149 (65.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak metropolitan influenced zone</td>
<td>88 (40.0)</td>
<td>132 (60.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No metropolitan influenced zone</td>
<td>8 (36.4)</td>
<td>14 (63.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health authority</strong></td>
<td></td>
<td></td>
<td></td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Interior Health</td>
<td>259 (42.8)</td>
<td>346 (57.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraser Health</td>
<td>754 (37.0)</td>
<td>1284 (63.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vancouver Coastal Health</td>
<td>1443 (31.2)</td>
<td>3180 (68.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Island Health</td>
<td>387 (37.6)</td>
<td>643 (62.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern Health</td>
<td>150 (49.7)</td>
<td>152 (50.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: IQR = interquartile range.  
*Except where noted otherwise. Percentages of row totals.  
†χ² for categoric variables, Kruskal–Wallis test for continuous variables.  
‡Data on ethnicity and injection drug use were collected from Drug Treatment Program enrolment forms completed by enrolling physicians.  
§Owing to increasingly recognized concerns about stigmatization with presenting disease outcomes by ethnicity in Canada and the limitations in existing data-stewardship policies for different groups, only data on the ethnic composition of the total study population are included.
model, living in the Northern Health region was associated with the lowest odds of virologic suppression.

Our findings have implications for both HIV care delivery and understanding limitations of rural health research at a population level. Appendix 1, Supplemental Figures S1 and S2 depict how different definitions of rurality capture different distributions of BC communities. Interestingly, when rurality was defined by Statistical Area Classification type, the most rural areas were not associated with poor virologic suppression, which suggests that the relation between rurality and virologic suppression is complex; however, small sample sizes may limit interpretation of this finding. A nonordinal relation with Statistical Area Classification type has also been observed for mortality. The strongest geographic predictor of virologic suppression appears to be health authority boundary, with poorer virologic suppression in the Northern Health region. Canada’s geography is unique, with a large area, but urban centres and most of the population are concentrated in southern areas. When making policy decisions around health care delivery, it is important to note that north–south discrepancies may be more important than rural–urban ones.

These findings highlight several methodologic considerations when including rurality in statistical analyses. First, researchers need to consider carefully why they are including rurality in their analysis and the mechanism by which rurality influences the outcome of interest. They should also consider whether rurality affects the health outcome directly or whether it is being used as a surrogate for a combination of unmeasured variables or variables that are difficult to measure. Accordingly, a definition of rurality should be selected that best matches the proposed mechanism of effect. For example, if community size and commuting patterns are most important, Statistical Area Classification is an appropriate definition. Second, researchers and health policy-makers must recognize that there is heterogeneity among rural areas and that generalized findings about rurality may not lead to locally responsive interventions. Researchers focusing on rural health outcomes can consider geographic information systems tools to present data, as they can visually show geographic differences that may be more useful to policy-makers.

Limitations

Our study is specific to the geography of Canada, so the findings may not be generalizable to other countries. We did not use all possible definitions of rurality. We looked only at 1 outcome for 1 disease, and patterns of health care access, disease incidence and factors that modify disease progression may differ for other chronic diseases. Finally, because people often change their place of residence, out-of-date geographic data may be a source of misclassification error.
Conclusion

We identified geographic discrepancies in HIV virologic suppression in BC, with poorer outcomes in northern and nonmetropolitan areas of the province. This finding highlights the need to improve access to HIV services outside of major urban centres, as well as to better understand factors contributing to rural health outcomes that may not be measured in population-level data sets. It also calls on researchers to think critically about how geographic variables are included in analyses and to consider paradigms other than just rural–urban as contributing to geographic health disparities in Canada. Additional studies using a rich collection of socioeconomic variables and data on use of health care services are needed to elucidate whether Rurality predicts poorer health outcomes or whether it is acting as a surrogate for other variables. More studies are also needed to better understand north–south differences in health outcomes in Canada.

References


Affiliations: Faculty of Medicine (Jaworsky), University of British Columbia, Vancouver, BC; Institute for Health Policy, Management and Evaluation (Jaworsky, Loutfy, Bayoumi) and Department of Medicine (Loutfy, Bayoumi, Richardson, Kuper), University of Toronto, Toronto, Ont.; Northern Medical Program (Jaworsky), University of Northern British Columbia, Prince George, BC; BC Centre for Excellence in HIV/AIDS (Jaworsky, Lu, Ye, Bratu, Sereda, Hogg), Vancouver, BC; Department of Medicine (Loutfy, Richardson), Women’s College Hospital; Division of General Internal Medicine (Bayoumi) and MAP Centre for Urban Health Solutions (Bayoumi), Li Ka Shing Knowledge Institute, St. Michael’s Hospital; University Health Network (Richardson); The Wilson Centre (Richardson, Kuper); Toronto, Ont.; Faculty of Health Sciences (Hogg), Simon Fraser University, Burnaby, BC

Contributors: Mona Loutfy, Ahmed Bayoumi, Lisa Richardson, Ayelet Kuper, Monica Ye, Paul Sereda and Robert Hogg supervised the project. Denise Jaworsky, Mona Loutfy, Ahmed Bayoumi, Lisa Richardson, Ayelet Kuper and Robert Hogg contributed to the study conceptualization and design. Denise Jaworsky and Michelle Lu analyzed the data. Denise Jaworsky interpreted the data, and the interpretation was reviewed by Mona Loutfy, Ahmed Bayoumi, Lisa Richardson and Ayelet Kuper. Denise Jaworsky drafted the manuscript, and Mona Loutfy, Ahmed Bayoumi, Lisa Richardson, Ayelet Kuper and Robert Hogg revised it critically for important intellectual content. All of the authors approved the final version to be published and agreed to be accountable for all aspects of the work.

Funding: The Comparative Outcomes and Service Utilization Trends study is funded by Operating Grant 130419 from the Canadian Institutes of Health Research (CIHR) and a Foundation Award (143342) to Robert Hogg. It is supported by the BC Centre for Excellence in HIV/AIDS. Denise Jaworsky received funding through a CIHR postdoctoral fellowship. Ahmed Bayoumi was supported by the Fondation Baxter & Alma Ricard Chair in Inner City Health at Unity Health Toronto and the University of Toronto.

Data sharing: The British Columbia Centre for Excellence in HIV/AIDS (BC-CfE) is prohibited from making this data set available publicly owing to prohibitions in the information-sharing agreement under which the data stewards provided the data to the BC-CfE. The underlying analytical codes are available from the authors on request.

Acknowledgements: The authors thank the Comparative Outcomes and Service Utilization Trends (COAST) study participants, the British Columbia Centre for Excellence in HIV/AIDS, the British Columbia Ministry of Health, the BC Vital Statistics Agency and the institutional data stewards for granting access to the data, and Population Data BC for facilitating the data linkage process. They also thank Simran Jawanda for her assistance in preparing the final manuscript.

Disclaimer: All inferences, opinions and conclusions drawn in this manuscript are those of the authors and do not reflect the opinions or policies of the data stewards or the funders.

Supplemental information: For reviewer comments and the original submission of this manuscript, please see www.cmajopen.ca/content/8/4/E643/suppl/DC1.