

Catalogue no. 82-622-X
ISSN 1915-5190
ISBN 978-0-660-72000-5

Working Paper

Health Region Peer Groups – Working paper, 2023

Release date: June 11, 2024



How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "[Standards of service to the public](#)."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

© His Majesty the King in Right of Canada, as represented by the Minister of Industry, 2024

Use of this publication is governed by the Statistics Canada [Open Licence Agreement](#).

An [HTML version](#) is also available.

Cette publication est aussi disponible en français.

Table of contents

Purpose	4
1 Introduction	4
2 Data	4
3 Methodology	5
3.1 Number of Clusters	5
4 Results	5
4.1 Standardization of Variables.....	5
4.2 Creation of Peer groups	6
4.3 Collapsing Small Clusters	6
4.4 Ontario Health Regions (OHR).....	7
5 Discussion	7
5.1 Strongest Predictors.....	7
5.2 Principal Component Analysis	8
5.3 Peer Group Description.....	8
5.4 Geographic Limitation	9
5.5 Geographic Representation of Final Peer Groups	9
6 Peer Groups in Action	10
6.1 Example: Heart Disease	11
7 Summary	12
8 References	13
Appendices	14
Appendix A Variable Definitions	14
Appendix B Descriptive statistics for final peer groups	15
Appendix C Descriptive summary of final peer groups.....	16
Appendix D Health Region Peer Groups.....	17

Health Region Peer Groups – Working paper, 2023

Purpose

The purpose of this document is to define the concept of peer groups, to give an overview of how they are created and to demonstrate their usefulness. This paper presents the classification of the 2023 peer groups.

1 Introduction

The launch of the Canadian Community Health Survey (CCHS) in 2000, combined with the expansion of existing data products at the health region level, prompted the desire for a method to compare regions with similar socioeconomic determinants of health. The reasoning behind developing such a method lies in the possibility of comparing regions by measures of health status once the effects of various social and economic characteristics known to influence health have been accounted for. This method enables the comparison of the relative effectiveness of health promotion and prevention activities across regions. Thus, the health regions have been grouped based on similar socioeconomic characteristics using a clustering technique, and these groups are referred to as ‘peer groups.’

Development of the criteria used to define peer groups required careful consideration of their intended use. The requirement that peer groups be used as a method for comparing health-related issues ultimately eliminated all variables directly describing health as potential candidates in the creation of the groups. Further, it was essential that all variables used be reliable and available for all health regions. As well, the need for objectivity required that peer groups be developed using empirical techniques. Finally, the need for simplified and relevant comparisons also required that peer groups consist of approximately 5 to 10 health regions per group. In the application of the above parameters, several limiting factors arose requiring some modifications. All criteria were adhered to the extent possible, and any deviations are thoroughly explained throughout this document.

The original 2000 Peer Group Classification was released in 2002 and was based on the 1996 Census information as well as the health region boundaries as defined by the provinces and territories in 2000. To remain current with respect to data availability and the health region boundary changes, it is necessary to update the peer group classification over time. These updates have occurred through the 2003, 2007, 2014 and 2018 Peer Group Classification. The latest update to the peer groups is based on the 2021 Census data and the health region boundaries as of September 2023. The final result of this classification was the creation of nine peer groups, representing all health regions across Canada.

This document provides an overview of the peer group creation. It presents the 2023 Peer Group Classification and compares the results with past peer groups. Finally, an example illustrates the use of peer groups in analyzing health-related issues.

2 Data

Typically, a set of 23 variables describing the socioeconomic and sociodemographic determinants of health within the health regions across Canada are used in the clustering algorithm to produce the peer groups. These variables encompass various subjects including demographic structure, social and economic status, ethnicity, Indigenous status, housing, urbanization, income inequality and labour market conditions. Note that health-related variables were deliberately not used in the creation of the peer groups.

There have been some modifications made over time; however, the majority of the variables have remained consistent since the creation of the 2000 Peer Group Classification. The 2023 peer group classification used the same 23 variables as those used in the creation of the 2018 peer groups. These variables are based on data from the 2021 Census. The variables used for the analysis, along with their respective sources, are outlined in Appendix A.

3 Methodology

A nonhierarchical cluster analysis was chosen as the method to create the peer groups. In essence, cluster analysis aims to assign observations to groups (clusters) based on a measure of their distance from each other so that observations within each group are similar to one another concerning the variables or attributes of interest. In other words, the goal is to group the observations into homogeneous and distinct clusters. Non-hierarchical algorithms seek to partition a set of observations into a predefined set of disjointed groups using a specified optimization criterion. This approach was deemed most suitable to meet the original objectives of the peer group project, mainly to use an empirical technique to create a predefined number of peer groups, each containing approximately 5 to 10 health regions.

The peer groups were created in SAS using the FASTCLUS procedure. This procedure uses a k -means algorithm to assign observations to a predefined set of k clusters. A description of k -means clustering, and several variants of the method can be found in Johnson and Wicheren (2002). The basic steps for placing observations into k clusters are as follows:

1. Select k observations as cluster seeds (the initial centres of the clusters).
2. Assign observations to the nearest cluster seed. After all observations are assigned, cluster seeds are replaced by their respective cluster means. This step is repeated until the change in cluster seeds becomes negligible or reaches zero.
3. Form final clusters by assigning each observation to its nearest cluster seed.

Complete details of the FASTCLUS procedure can be found in the SAS® 9.4 and SAS® Viya® 3.5 Programming Documentation.

3.1 Number of Clusters

One of the major challenges with cluster analysis is selecting the appropriate number of initial clusters. Several criteria have been suggested (Everitt et al., 2001) which generally involve the optimization of one or more test statistics. From a practical perspective, it is generally left up to the analyst to determine the number that best suits a given need. For the 2023 Peer Group classification, a maximum of 15 clusters was chosen. This would give an average number of 7 health regions to each peer group,¹ which is in line with the study objectives. The maximum number of clusters used in 2018 was 16. The number of health regions in 2023 is also lower than it was in 2018.

4 Results

4.1 Standardization of Variables

Variables measured on different scales, or on a common scale with differing variances, are often standardized to mitigate the effect of these differences between the variables. For this exercise, all 23 socioeconomic variables were standardized (mean 0, variance 1) prior to performing the cluster analysis.

Some variables contained missing or zero values to indicate that this information was not available in certain health regions. The proportion of low-income individuals in private households (LOWPOP) and the proportion of low-income children (LOWKIDS) contained missing values, because the Census does not derive low-income data for the three territories and Indian reserves. The regions containing missing values for these variables were “Région des Terres-Cries-de-la-Baie-James” (2418) and the territories. All missing values were set to zero for the analysis, before standardizing the variables.

There was another variable on the Census file that had a value of zero for some health regions. A value of zero for the MIZ (Metropolitan Influenced Zone) variable simply means that there is not a large metropolitan area within the health region. Therefore, a value of zero on the data file is not necessarily an indication that the variable could not be derived.

1. Note that peer group and cluster are used interchangeably to refer to the classification of health regions into groups with similar socioeconomic characteristic.

4.2 Creation of Peer groups

To establish a starting point, the clustering algorithm was instructed to group the health regions into 15 clusters. Five of the resulting clusters contained only one health region. This suggests that using 15 clusters may be excessive, as the primary aim of creating peer groups is to facilitate comparisons among similar health regions. The cluster analysis was rerun with a reduced number of cluster seeds.

The results of the final cluster analysis using PROC FASTCLUS can be seen in Table 4.2.1. The table shows the number of health regions contained in each peer group, as well as several statistics related to the clusters. The root mean square standard deviation is a measure of the variability in the data points around the cluster centre. The radius displays the largest Euclidean distance from the cluster centre to any observation within the cluster. The nearest cluster refers to the closest peer group in terms of Euclidean distance. Finally, the last column of the table displays the distance between the current cluster centre and that of its closest neighbour. For each of these statistics, the cluster centre is the point having coordinates that are the means of all the observations in the cluster. Euclidean distance is a statistical measure of distance between two points.

Table 4.2.1
Results of final cluster analysis using PROC FASTCLUS

Cluster	Frequency	Root Mean Square Standard Deviation	Radius	Nearest Cluster	Distance Between Cluster Centres
A	5	0.97	5.01	E	9.14
B	12	0.54	3.35	I	4.64
C	4	0.93	4.61	J	6.37
D	2	0.75	2.54	E	5.03
E	18	0.52	4.37	I	2.84
F	11	0.51	3.6	I	3.24
G	11	0.52	3.76	H	3.22
H	6	0.58	3.0	G	3.22
I	29	0.44	3.13	E	2.84
J	7	0.55	2.97	H	3.46

Source: Results of the health regions clustering analysis conducted using 23 indicators from the 2021 Census.

There were two clusters that contained the majority of the health regions (E and I); both had low root mean square standard deviation and the lowest distance between cluster centres. The fact that they were nearest neighbours also indicates that all the regions in these clusters were very similar. Therefore, although these clusters were large and they go against the objective of having approximately 5 to 10 regions per peer group, there did not appear to be a valid reason to split them into smaller groups.

4.3 Collapsing Small Clusters

The results in Table 4.2.1 represent clusters that are roughly evenly distributed and have minimal within cluster variance based on the parameters used by the clustering algorithm. The results indicate the formation of 10 clusters, varying in size from 2 to 29 health regions. However, having a cluster with fewer than five regions is not practical as it limits options for comparison. To enhance comparability, clusters with less than five members were combined with their nearest neighbour. The exception was cluster C (Montréal, Toronto, Vancouver and Richmond). Cluster C was not combined with another cluster since these health regions tend to be very different than other regions across the country.

There was one cluster that was joined with its closest neighbour. Cluster D (health regions 6001 and 6101) was combined with its nearest neighbour cluster E. The collapsing of clusters D and E produced a cluster with 20 health regions, so no additional collapsing was required. This combined cluster was labelled Cluster E. The result of collapsing the smaller clusters was that the 10 peer groups produced from the final cluster analysis using the FASTCLUS procedure and presented in Table 4.2.1 were reduced to 9 groups. To maintain continuity in the alphabetical nomenclature of peer groups, cluster J was renamed cluster D. A list of Health regions categorized by the final peer groups can be found in Appendix D.

4.4 Ontario Health Regions (OHR)

There are two levels of geography in Ontario: there are 6 Ontario Health Regions (OHR) and 34 Public Health Units (PHU). Due to the relationship between the two levels of geography it was possible to incorporate both into the peer group classification. The information at the PHU level was used to create the peer groups. At the final stage in the cluster analysis, the OHR level geography was added to the existing clusters. The OHR did not have an impact on the placement of the other health regions into the final peer groups. In an analysis involving the peer groups, only one level of geography in Ontario should be used.

Table 4.4.1
Peer groups for the OHR in Ontario

OHR	Name	Peer Group
3501	West	I
3502	Central	H
3503	Toronto	C
3504	East	I
3505	North East	B
3506	North West	E

5 Discussion

5.1 Strongest Predictors

To determine which variables played a key role in defining the health region peer groups, the final clusters were run against all 23 variables in a stepwise discriminant analysis. Partial R² statistics for entry and removal were set at 0.15. Any variable which had an R² value of 0.5 or higher when regressed against a variable already in the model was removed from the analysis. Overall, five variables appeared to be the most important predictors. Table 5.1.1 displays a summary of the results.

Table 5.1.1
Stepwise discriminant analysis of final health region groupings on the 23 variables

Step	Variable	Partial R-SQ	R-SQ Variables in Model
1 (Indigenous Identity)	INDIG_RATE	0.886	...
	Removed
	None
2 (Population Density)	POPDEN	0.8243	...
	Removed
	LNEPRNT	...	0.7424
	OWNDWL	...	0.5356
	POP20	...	0.6529
	POSTSEC	...	0.5966
3 (Immigrants 2011-2021)	IMMPER	0.7203	...
	Removed
	LOWPOP	...	0.5401
4 (Average Value of Dwelling)	AVGDWL	0.5116	...
	Removed
	GOVTRAN	...	0.5589
	HOUAFF	...	0.6937
	LOWKIDS	...	0.5105
	MEDINC	...	0.5838
	MEDSHR	...	0.5961
	MIZ	...	0.5225
	POP21	...	0.5622
	POP65	...	0.7207
	VISMIN	...	0.8567

Table 5.1.1
Stepwise discriminant analysis of final health region groupings on the 23 variables

Step	Variable	Partial R-SQ	R-SQ Variables in Model
5 (Long-term Unemployment Rate)	LTUNEMP	0.4993	...
	Removed
	GROWTH	...	0.615
	MIGMOB	...	0.5858

... not applicable

Source: Results of the discriminant analysis conducted on the 23 indicators from the 2021 Census used for clustering.

The strongest predictors of the final peer groups were Indigenous identity and population density. Four variables (LTUNEMP, OWNDWL, POP20, POSTSEC) were removed from the analysis when regressed against Indigenous identity, whereas LOWPOP was removed from the analysis when regressed against population density.

5.2 Principal Component Analysis

Principal component analysis is a multivariate technique which aims to reduce the number of variables in the data to a few factors called principal components. Principal components are linear combinations of the original variables and are uncorrelated. They are derived in decreasing order of importance, so that as much of the total variance in the data can be explained in as few factors as possible. Therefore, the first principal component holds the highest importance as it accounts for the largest proportion of the total variance present in the dataset.

A principal component analysis was performed on the 23 socioeconomic variables used in the cluster analysis. The first two principal components accounted for just over 57% of the total variability. The first principal component appears to represent factors associated with “urbanicity,” including housing affordability, proportion of visible minorities, proportion of immigrants, average dwelling value, and population living in census metropolitan areas. The second principal component seems indicative of characteristics related to family profile, including the proportion of the population aged 65 and over, proportion of lone-parent families, proportion of the population aged 0 to 19, proportion of owner-occupied dwellings, and proportion of Indigenous population. As for the third principal component, it could be interpreted as reflecting income inequality, as evidenced by such variables as median household income, proportion of all income that came from government transfers, proportion of low-income children, and proportion of low-income individuals in private households. The first six principal components accounted for over 88% of the total variability in the data, showing that 23 variables can be effectively reduced to six factors with minimal loss of information. These results are similar to the previous peer group classification, which indicates that the variables which drive the analysis are remaining fairly consistent over time.

5.3 Peer Group Description

The five key variables determined by the stepwise discriminant analysis were used to represent each of the clusters. The mean values of these five variables for each peer group can be found in Appendix B. For each of the five variables, several percentiles were calculated and used to classify the peer groups. Values were classified based on the following ranges.

Very High: $X > 85\text{th percentile}$

High: $65\text{th percentile} < X \leq 85\text{th percentile}$

Medium: $35\text{th percentile} < X \leq 65\text{th percentile}$

Low: $15\text{th percentile} < X \leq 35\text{th percentile}$

Very Low: $X \leq 15\text{th percentile}$

The results from this classification can be found in Table 5.3.1. While the methodology is simplistic as a descriptive tool, it effectively distinguishes the characteristics of one peer group from another. As shown in the table below, no two peer groups share the same category for all five variables. For example, peer group C (comprising Montréal, Toronto, Vancouver and Richmond) is the only group characterized by very low proportion of individuals identifying

as Indigenous, very high population density, very high proportion of immigrants who arrived between 2011 and 2021, and a very high average value of dwellings.

Table 5.3.1
Final peer grouping descriptions based on five factors resulting from the stepwise discriminant analysis

Cluster	Indigenous Identity	Population Density	Immigrants 2011-2021	Average Value of Dwelling	Long-term Unemployment Rate
A	Very High	Very Low	Low	Medium	Very High
B	Medium	Medium	Very Low	Very Low	Very High
C	Very Low	Very High	Very High	Very High	Medium
D	Low	Very High	Very High	Very High	High
E	High	Low	Medium	Low	Medium
F	Very Low	Medium	Low	Low	Very Low
G	Low	High	High	Medium	Medium
H	Medium	High	High	Very High	Low
I	Medium	Medium	Medium	High	Medium

Source: Summary of the discriminant analysis results conducted on the 23 indicators from the 2021 Census used for clustering.

The results from this classification were used to derive a written summary of the nine peer groups based on the five key variables from the discriminant analysis. This summary is presented in Appendix C.

5.4 Geographic Limitation

Each province and territory defines the geographic boundaries for a health region based on administrative preference and these boundary definitions change over time. Health regions can be strictly urban or rural or some combination of the two. There may be considerable variability within health regions regarding health measures due to the lack of geographic homogeneity. This variability should be taken into consideration when making inferences about a certain region. For instance, even though the health indicators in Vancouver compare favourably with the national averages, this should not be interpreted as meaning that the residents of the downtown core in Vancouver have better than average health. This lack of homogeneity in defining health region boundaries complicates the process of assigning health regions to peer groups. This variability can significantly affect how well a specific variable represents the entire region, and in some cases, important defining factors may be overlooked.

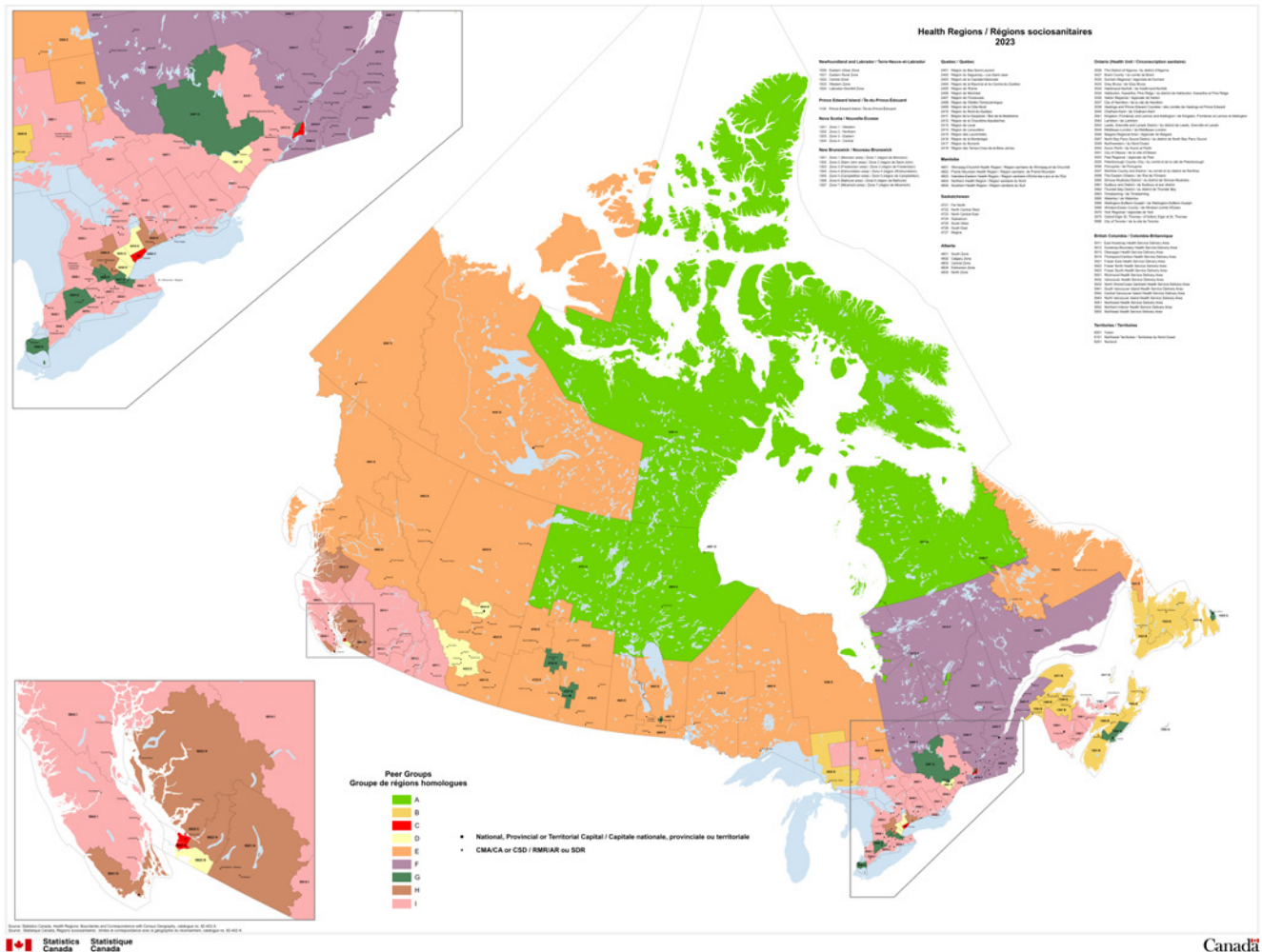
It should also be noted that there may be considerable variability amongst the health regions within a peer group in regard to the socioeconomic factors used in the cluster analysis. This should be considered when comparing regions within a certain peer group. This variability is apparent among the 2023 peer groups listed in Appendix B, showcasing the diversity across the five key variables identified through stepwise discriminant analysis.

5.5 Geographic Representation of Final Peer Groups

The map below provides a clear visual representation of the geographic clustering of the health regions into the final 9 peer groups. Montréal, Toronto, Vancouver, and Richmond constitute the smallest cluster due to their significant differences in population size and diversity compared to the other health regions, making them unsuitable for combination with any other peer group.

Clusters of health regions have clearly formed, largely due to shared characteristics shaped by their geographical location within Canada. For instance, the northern regions have clustered together based on the Indigenous composition of their communities and the low population density.

Map 1
Health Regions and Peer Groups in Canada, 2023



A larger version of the map is [available](#).

6 Peer Groups in Action

The purpose of this section is to illustrate the usefulness of the peer groups. There are two valuable, yet distinct, analyses possible with peer groups: comparing health-related indicators between and within peer groups. Since peer groups are formed based on regions that have similar socioeconomic characteristics, differences between peer groups are expected to emerge. Peer groups exhibiting better socioeconomic status indicators are likely to have better health status measures. Additionally, estimates from a single peer group can be compared with national averages to assess the overall performance of the group of regions. The second analysis possible, perhaps of greater relevance, is the comparison of health regions within a peer group. Once the effects of the various socioeconomic characteristics known to influence health status have been removed, a more meaningful comparison of regions based on health status measures becomes possible.

The example provided in Section 6.1 is a simple illustration of how and when peer groups can be used. The example uses 2023 Peer Group Classification and 2019–2020 Canadian Community Health Survey (CCHS) data. A more comprehensive analysis involving the peer groups can be found in the paper “The Health of Canada’s Communities” authored by Margot Shields and Stéphane Tremblay of Statistics Canada (2002).

6.1 Example: Heart Disease

This example examines the prevalence of heart disease in the population 18 years of age and over in the different regions across the country. Every CCHS respondent is asked about their heart disease status. The national prevalence of heart disease among the adult population in 2019-2020 was 5.0%. The rate of missing data for this health indicator is less than 0.5%. In this example, the missing values have been excluded.

The prevalence of heart disease in each peer group is shown in Table 6.1.1, along with a description of each peer group. The prevalence of heart disease in Peer Group D is one percentage point lower than the national average. It is also 1.6 percentage points lower than the rate of heart disease in Peer Group E. Both of these differences are statistically significant (p -value<0.01). Peer Group D is composed of large cities and suburbs in Ontario, Alberta and British Columbia, characterized by a very high population density. This group exhibits a low smoking rate (11.7%), a low heavy drinking rate (15.8%) and an above-average exercise rate (72.3%). On the other hand, Peer Group E comprises mainly Northern regions in Ontario and British Columbia and rural regions in the Prairies with a low population density. This group has a higher smoking rate (19.7%), a higher heavy drinking rate (21.6%) and a lower physical activity rate (66.0%). Except for smoking, the differences in the rates of these risk factors between Peer Groups D and E are statistically significant (p -value<0.01).

Table 6.1.1
Prevalence of Heart Disease by Peer Group

Peer Group	Number of Health Regions	Principal Characteristics	Heart Disease Prevalence
A	5	<ul style="list-style-type: none"> Northern and remote regions with very low population density Very high long-term unemployment rate Very high proportion of Indigenous population 	3.9% [2.0%, 5.8%]
B	12	<ul style="list-style-type: none"> Mainly rural Eastern regions with low population density Very low proportion of immigrants who arrived between 2011 and 2021 Very low average dwelling value Very high long-term unemployment rate 	8.4% [7.7%, 9.1%]
C	4	<ul style="list-style-type: none"> Largest metro centres (Toronto, Montréal, Vancouver/Richmond) Very low proportion of Indigenous population Very high population density Very high proportion of immigrants who arrived between 2011 and 2021 Very high average value of dwelling 	3.9% [3.3%, 4.5%]
D	7	<ul style="list-style-type: none"> Large cities and suburbs in Ontario, Alberta and British Columbia Very high population density Very high proportion of immigrants who arrived between 2011 and 2021 Very high average value of dwelling High long-term unemployment rate 	4% [3.6%, 4.5%]
E	20	<ul style="list-style-type: none"> Mainly Northern regions in Ontario and British Columbia, rural regions in the Prairies, and Yukon, and Northwest Territories High proportion of Indigenous population Low population density 	5.6% [5.2%, 6.1%]
F	11	<ul style="list-style-type: none"> Regions in Québec outside of Montréal Very low proportion of Indigenous population Low proportion of immigrants who arrived between 2011 and 2021 Very low long-term unemployment rate 	6.1% [5.6%, 6.7%]
G	11	<ul style="list-style-type: none"> Mainly urban centres with high population density High proportion of immigrants who arrived between 2011 and 2021 Medium average dwelling value 	4.9% [4.4%, 5.3%]
H	6	<ul style="list-style-type: none"> Regions around the Toronto and Vancouver areas High population density High proportion of immigrants who arrived between 2011 and 2021 Very high average dwelling value 	3.9% [3.4%, 4.5%]
I	29	<ul style="list-style-type: none"> Sparsely populated urban-rural mix from coast to coast Medium population density High average dwelling value 	6.2% [5.8%, 6.6%]

Note: Values in brackets represent the lower and upper limits of the 95% confidence interval.

Sources: Statistics Canada, Canadian Community Health Survey (CCHS), 2019 and 2020.

Peer Group D comprises seven health regions. Table 6.1.2 shows the prevalence of heart disease in each of these regions. Six out of the seven regions have a prevalence equal to or below the national prevalence of 5.0%. The highest prevalence of heart disease is 5.1%, it is observed in health region 3551. On the other hand, the lowest prevalence of heart disease is 3.3% is found in health region 4832.

Table 6.1.2
Prevalence of heart disease in Health Regions belonging to Peer Group D

Health Region	Name	Heart Disease Prevalence
3536	Halton Regional Health Unit	4.3% [2.7%, 5.9%]
3551	City of Ottawa Health Unit	5.1% [3.6%, 6.5%]
3553	Peel Regional Health Unit	3.8% [2.7%, 4.8%]
3570	York Regional Health Unit	4.8% [3.5%, 6.2%]
4832	Calgary Zone	3.3% [2.4%, 4.1%]
4834	Edmonton Zone	3.6% [2.8%, 4.4%]
5923	Fraser South Health Service Delivery Area	4.4% [3.2%, 5.6%]

Note: Values in brackets represent the lower and upper limits of the 95% confidence interval.

Sources: Statistics Canada, Canadian Community Health Survey (CCHS), 2019 and 2020.

Note that the heart disease prevalences presented in the tables 6.1.1 and 6.1.2 above can all be published without caution, as a sufficient number of respondents were included in their calculations. The confidence interval can be used to assess the reliability of the estimate itself.

For peer groups containing more remote health regions, conducting the same analysis may not be feasible due to the small number of respondents. In such cases, the results are typically published at the province level to obtain a larger sample size and more reliable estimates. The peer groups offer an alternative to the provinces in these situations.

7 Summary

As a result of health region boundary changes as of September 2023, and the availability of 2021 Census data, it was necessary to update the 2018 Peer group classification. In keeping with the original working paper, the goal was to produce a classification which would cluster health regions with similar social and economic health determinants into peer groups. Twenty-three variables covering a wide range of social, economic and demographic areas were used to cluster the health regions.

Starting with an initial set of 15 clusters and ensuring that each cluster contained at least two health regions, the results indicated that the regions naturally organized themselves into 10 distinct peer groups. Peer groups with fewer than five health regions were combined with their nearest neighbour. This was done to ensure an adequate number of health regions within a peer group for comparison purposes. Cluster C, consisting of Montréal, Toronto, Vancouver and Richmond was not merged with another cluster, as these health regions share more similarities among themselves than with others. The final result comprised 9 peer groups ranging in size from 4 to 29 (not including the Ontario Health Regions (OHR) in Ontario).

Stepwise discriminant analysis was used to determine which variables had the most influence on the final peer groupings. The five most important variables were Indigenous identity, population density, proportion of immigrants who arrived between 2011 and 2021, average value of dwelling, and long-term unemployment rate. Each peer group is characterized by at least one distinguishing factor in terms of these five variables.

Peer groups are valuable for analyzing health-related indicators because, after accounting for the effects of various social and economic characteristics known to influence health status, a more meaningful comparison of regions becomes feasible. Health indicators can be compared both between and within peer groups. Additionally, peer groups serve as an alternative to provinces when the results of an analysis cannot be presented at the health region level due to insufficient sample size or high sampling variability.

8 References

Andberg, M. R. (1973). *Cluster Analysis for Applications*. New York: Academic Press.

Everitt, B. S., Landau S., Leese M. and Stahl, D. (2001). *Cluster Analysis*, 5th Edition, John Wiley and Sons, Ltd: Chichester, UK.

Johnson, R. and Wicheren, D. (2002). *Applied Multivariate Statistical Analysis*, Prentice Hall

Sarafin, C. (2009). 2007 Health Region Peer Groups Methodology Guide, *Internal document, Health Indicators*, Statistics Canada.

SAS Institute Inc., [SAS/STAT® User's Guide](#).

Shields, M. and Tremblay, S. (2002). *The Health of Canada's Communities*, Statistics Canada.

Wannell B. (2009). 2007 Health Region Peer Groups, *Internal document, Health Indicators*, Statistics Canada

Appendices

Appendix A

Variable Definitions

Variables used in the creation of the 2023 Peer Groups are from Census 2021

Variable	Description
AVGDWL	Average value of dwelling -owner-occupied, non-farm, non-reserve (Canadian dollars)
EMP	Employment rate (persons aged 25 to 54)
GOVTRAN	Government transfer income in 2020, as a proportion of total income (percent)
GROWTH	Growth rate (% change in regions population between 2016 and 2021)
HOUAFF	Households spending 30% or more of household income on shelter, proportion of total shelter-cost households
IMMPER	Immigrants who arrived between 2011 and 2021, proportion of total population (percent)
INDIG_RATE	Indigenous identity population, proportion of total population (percent)
LNEPRNT	Lone-parent families, proportion of census families (percent)
LOWKIDS	Prevalence of persons aged 17 years and under living in low-income economic families before tax in 2020 (percent)
LOWPOP	Prevalence of low income before tax in 2020 for persons in private households (percent)
LTUNEMP	Long-term unemployment rate, labour force aged 15 and over
MEDINC	Median household income
MEDSHR	Income share held by households whose incomes fall below the median household income in 2020 (percent)
MIGMOB	5-year internal migrants, proportion of population aged 5 years and over (percent)
MIZ	Population living within a Census Metropolitan Area, a Census Agglomeration or a strong Census Metropolitan Area and Census Agglomeration Influenced Zone (percent)
OWNDWL	Owner-occupied private non-farm, non-band, non-reserve dwellings (percent)
POP20	Population aged 0 to 19 years, proportion of total population
POP21	2021 population (based on population and dwelling counts not randomly rounded but adjusted for areas with a pop < 20)
POP65	Population aged 65 years and over, proportion of total population
POPDEN	Population density (population per square kilometer) (number)
POSTSEC	Post-secondary graduates aged 25 to 54, proportion of population aged 25 to 54 (percent)
UNEMP	Unemployment rate 15 years and over
VISMIN	Visible minority population, proportion of total population (percent)

Appendix B

Descriptive statistics for final peer groups

Cluster		Indigenous Identity	Population Density	Immigrants 2011-2021	Average Value of Dwelling	Long-term unemployment Rate
A	N	5	5	5	5	5
	MIN	73.1	0.02	0.4	185,000	8
	MAX	95.5	3.04	2.2	470,000	19.2
	Mean	86.4	0.68	1	325,320	13.5
	St. Dev.	8.5	1.31	0.7	124,149	4.6
B	N	12	12	12	12	12
	MIN	2.1	2.14	0.3	135,000	9.7
	MAX	27.8	16.08	1.2	266,800	20.9
	Mean	8.8	6.76	0.8	192,916	14.9
	St. Dev.	6.9	4.52	0.3	38538	3.5
C	N	4	4	4	4	4
	MIN	0.7	1629	10.8	638,000	9
	MAX	2.3	5186	13.3	1,728,000	13.9
	Mean	1.2	3816	11.9	1,168,750	11
	St. Dev.	0.8	1536	1.2	445708	2.1
D	N	7	7	7	7	7
	MIN	0.5	42.7	7.9	434,000	8
	MAX	6.3	1163.2	13.4	1,230,000	13.5
	Mean	2.5	574	10.6	873,929	11.3
	St. Dev.	2.1	426.9	1.7	321,866	1.8
E	N	20	20	20	20	20
	MIN	5.9	0.04	0.7	210,400	5.9
	MAX	49.6	7.6	5.6	488,800	16
	Mean	22.1	1.75	3.4	310,700	9
	St. Dev.	11.8	1.97	1.7	69,147	2.2
F	N	11	11	11	11	11
	MIN	1.4	0.05	0.4	170,600	4.5
	MAX	16.2	166.2	4.2	411,200	7.5
	Mean	5.1	31.4	1.5	257,727	6.1
	St. Dev.	5.5	47.9	1.2	75,980	0.9
G	N	11	11	11	11	11
	MIN	0.8	13.3	2.2	344,000	8.2
	MAX	12.3	1781	11.9	752,000	15.7
	Mean	5.1	420.6	7.4	476,764	10.6
	St. Dev.	4.1	567	2.9	149,382	2.2
H	N	6	6	6	6	6
	MIN	1.7	5.7	4.3	802,000	6.8
	MAX	7.5	312.8	10.7	1,496,000	12.8
	Mean	4	145.6	6.5	991,833	8.9
	St. Dev.	2.3	130.3	2.6	266,678	2.1
I	N	29	29	29	29	29
	MIN	1.7	2	0.6	222,600	7
	MAX	14.2	258	4.3	750,000	15.5
	Mean	5.78	36.1	2	490,131	10
	St. Dev.	3.5	49.7	1	144,174	1.8

Appendix C**Descriptive summary of final peer groups**

Peer Group	Number of Health Regions	Percent of Canadian Population	Principal Characteristics
A	5	0.5%	<ul style="list-style-type: none"> Northern and remote regions with very low population density Very high proportion of Indigenous population Very high long-term unemployment rate
B	12	3.2%	<ul style="list-style-type: none"> Mainly rural Eastern regions with low population density Very low proportion of immigrants who arrived between 2011 and 2021 Very low average dwelling value Very high long-term unemployment rate
C	4	15.4%	<ul style="list-style-type: none"> Largest metro centres (Toronto, Montreal, Vancouver/Richmond) Very low proportion of Indigenous population Very high population density Very high proportion of immigrants who arrived between 2011 and 2021 Very high average dwelling value
D	7	22.1%	<ul style="list-style-type: none"> Large cities and suburbs in Ontario, Alberta and British Columbia Very high population density Very high proportion of immigrants who arrived between 2011 and 2021 Very high average dwelling value High long-term unemployment rate
E	20	8.0%	<ul style="list-style-type: none"> Mainly Northern regions in Ontario and British Columbia, rural regions in the Prairies, and Yukon, and Northwest Territories High proportion of Indigenous population Low population density
F	11	13.2%	<ul style="list-style-type: none"> Regions in Québec outside of Montréal Medium population density Very low proportion of Indigenous population Very low long-term unemployment rate Low proportion of immigrants who arrived between 2011 and 2021
G	11	13.4%	<ul style="list-style-type: none"> Mainly urban centres with high population density High proportion of immigrants who arrived between 2011 and 2021 Medium average dwelling value
H	6	7.4%	<ul style="list-style-type: none"> Regions around the Toronto and Vancouver areas High population density High proportion of immigrants who arrived between 2011 and 2021 Very high average dwelling value
I	29	16.9%	<ul style="list-style-type: none"> Sparsely populated urban-rural mix from coast to coast Medium population density High average dwelling value

Appendix D

Health Region Peer Groups

Peer group A

2417	Région du Nunavik
2418	Région des Terres-Cries-de-la-Baie-James
4604	Northern Health Region
4721	Far North
6201	Nunavut

Peer group B

1021	Eastern Rural Zone
1022	Central Zone
1023	Western Zone
1201	Zone 1 - Western
1202	Zone 2 - Northern
1203	Zone 3 - Eastern
1304	Zone 4 (Edmundston area)
1305	Zone 5 (Campbellton area)
1306	Zone 6 (Bathurst area)
1307	Zone 7 (Miramichi area)
2411	Région de la Gaspésie—Îles-de-la-Madeleine
3526	The District of Algoma Health Unit

Peer group C

2406	Région de Montréal
3595	City of Toronto Health Unit
5931	Richmond Health Service Delivery Area
5932	Vancouver Health Service Delivery Area

Peer group D

3536	Halton Regional Health Unit
3551	City of Ottawa Health Unit
3553	Peel Regional Health Unit
3570	York Regional Health Unit
4832	Calgary Zone
4834	Edmonton Zone
5923	Fraser South Health Service Delivery Area

Peer group E

1024	Labrador-Grenfell Zone
3549	Northwestern Health Unit
3556	Porcupine Health Unit
3562	Thunder Bay District Health Unit
3563	Timiskaming Health Unit
4602	Prairie Mountain Health Region
4603	Interlake-Eastern Health Region

Appendix D

Health Region Peer Groups

4605	Southern Health Region
4722	North Central West
4723	North Central East
4725	South West
4726	South East
4831	South Zone
4833	Central Zone
4835	North Zone
5951	Northwest Health Service Delivery Area
5952	Northern Interior Health Service Delivery Area
5953	Northeast Health Service Delivery Area
6001	Yukon
6101	Northwest Territories
Peer group F	
2401	Région du Bas-Saint-Laurent
2402	Région du Saguenay—Lac-Saint-Jean
2403	Région de la Capitale-Nationale
2404	Région de la Mauricie et du Centre-du-Québec
2405	Région de l'Estrie
2408	Région de l'Abitibi-Témiscamingue
2409	Région de la Côte-Nord
2410	Région du Nord-du-Québec
2412	Région de la Chaudière-Appalaches
2414	Région de Lanaudière
2416	Région de la Montérégie
Peer group G	
1020	Eastern Urban Zone
1204	Zone 4 - Central
2407	Région de l'Outaouais
2413	Région de Laval
3537	City of Hamilton Health Unit
3544	Middlesex-London Health Unit
3565	Waterloo Health Unit
3568	Windsor-Essex County Health Unit
4601	Winnipeg-Churchill Health Region
4724	Saskatoon
4727	Regina
Peer group H	
3530	Durham Regional Health Unit
3566	Wellington-Dufferin-Guelph Health Unit

Appendix D

Health Region Peer Groups

5921	Fraser East Health Service Delivery Area
5922	Fraser North Health Service Delivery Area
5933	North Shore/Coast Garibaldi Health Service Delivery Area
5941	South Vancouver Island Health Service Delivery Area
Peer group I	
1100	Prince Edward Island
1301	Zone 1 (Moncton area)
1302	Zone 2 (Saint John area)
1303	Zone 3 (Fredericton area)
2415	Région des Laurentides
3527	Brant County Health Unit
3533	Grey Bruce Health Unit
3534	Haldimand-Norfolk Health Unit
3535	Haliburton, Kawartha, Pine Ridge District Health Unit
3538	Hastings and Prince Edward Counties Health Unit
3540	Chatham-Kent Health Unit
3541	Kingston, Frontenac and Lennox and Addington Health Unit
3542	Lambton Health Unit
3543	Leeds, Grenville and Lanark District Health Unit
3546	Niagara Regional Area Health Unit
3547	North Bay Parry Sound District Health Unit
3550	Huron Perth Health Unit
3555	Peterborough County—City Health Unit
3557	Renfrew County and District Health Unit
3558	The Eastern Ontario Health Unit
3560	Simcoe Muskoka District Health Unit
3561	Sudbury and District Health Unit
3575	Oxford Elgin St. Thomas Health Unit
5911	East Kootenay Health Service Delivery Area
5912	Kootenay-Boundary Health Service Delivery Area
5913	Okanagan Health Service Delivery Area
5914	Thompson/Cariboo Health Service Delivery Area
5942	Central Vancouver Island Health Service Delivery Area
5943	North Vancouver Island Health Service Delivery Area