Catalogue no. 82-622-X
ISSN 1915-5190
ISBN 978-0-660-78595-0

Working Paper

Health Region Peer Groups: Working paper, 2024

Release date: October 2, 2025

Statistics Canada
Statistique Canada

Canada

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, the Agency has developed standards of service which its employees observe in serving its clients. 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.

Table of contents

Health Region Peer Groups: Working paper, 2024

Purpose

The purpose of this document is to define the concept of health region peer groups, provide an overview of their creation process, and demonstrate their practical value. This paper presents the classification of the 2024 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 need for a method to compare regions with similar socioeconomic determinants of health. The rationale behind developing such a method lies in the ability to compare regions by measures of health status after accounting for the effects of various social and economic factors known to influence health. This method enables the comparison of the relative effectiveness of health promotion and prevention activities across regions. To support meaningful comparison, health regions have been grouped into 'peer groups' based on similar socioeconomic characteristics using a clustering technique.

The development of the criteria used to define peer groups required careful consideration of their intended purpose. Since the primary goal was to enable comparisons of health-related issues, variables that directly described health outcomes were excluded from the grouping process. Additionally, all selected variables needed to be reliable and consistently available across all health regions. To ensure objectivity, empirical methods were used to develop peer groups. Finally, to facilitate simplified and relevant comparisons, peer groups were designed to consist of approximately 5 to 10 health regions per group. While applying these parameters, several constraints arose, requiring some adjustments. All criteria were followed as closely as possible, and any deviations are thoroughly explained throughout this document.

The original 2000 peer group classification was released in 2002 and was based on 1996 Census data, along with 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, 2018, and 2023 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 2024. This latest classification resulted in the creation of ten peer groups, encompassing all health regions across Canada.

Table 1.1
Health Region Peer Groups

Peer group A	
1020	Eastern Urban Zone
1204	Zone 4 - Central
2413	Région de Laval
3537	City of Hamilton Health Unit
3544	Middlesex-London Health Unit
3551	City of Ottawa Health Unit
3565	Waterloo Health Unit
3568	Windsor-Essex County Health Unit
4601	Winnipeg-Churchill Health Region
4724	Saskatoon Zone
4727	Regina Zone
4834	Edmonton Zone

Table 1.1
Health Region Peer Groups

Peer group B	
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
Peer group C	
1024	Labrador-Grenfell Zone
3549	Northwestern Health Unit
3556	Porcupine Health Unit
4722	North Central West Zone
4723	North Central East Zone
4833	Central Zone
4835	North Zone
5950	Northern Health Authority
6001	Yukon
6101	Northwest Territories
Peer group D	
2417	Région du Nunavik
2418	Région des Terres-Cries-de-la-Baie-James
4604	Northern Health Region
4721	Far North Zone
6201	Nunavut
Peer group E	
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

Table 1.1
Health Region Peer Groups

Peer group F	
3530	Durham Regional Health Unit
3536	Halton Regional Health Unit
3553	Peel Regional Health Unit
3570	York Regional Health Unit
4832	Calgary Zone
5920	Fraser Health Authority
5930	Vancouver Coastal Health Authority
Peer group G	
1302	Zone 2 (Saint John area)
1303	Zone 3 (Fredericton area)
3540	Chatham-Kent Health Unit
3542	Lambton Health Unit
3550	Huron Perth Health Unit
3558	The Eastern Ontario Health Unit
3561	Sudbury and District Health Unit
3562	Thunder Bay District Health Unit
3563	Timiskaming Health Unit
4602	Prairie Mountain Health Region
4603	Interlake-Eastern Health Region
4605	Southern Health Region
4725	South West Zone
4726	South East Zone
4831	South Zone
Peer group H	
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
3541	Kingston, Frontenac and Lennox and Addington Health Unit
3543	Leeds, Grenville and Lanark District Health Unit
3546	Niagara Regional Area Health Unit
3547	North Bay Parry Sound District Health Unit
3555	Peterborough County–City Health Unit
3557	Renfrew County and District Health Unit
3560	Simcoe Muskoka District Health Unit
3566	Wellington-Dufferin-Guelph Health Unit
3575	Oxford Elgin St. Thomas Health Unit

Table 1.1
Health Region Peer Groups

5910	Interior Health Authority
5940	Island Health Authority
Peer group I	
2406	Région de Montréal
3595	City of Toronto Health Unit
Peer group J	
1100	Prince Edward Island
1301	Zone 1 (Moncton area)
2407	Région de l'Outaouais
2414	Région de Lanaudière
2415	Région des Laurentides
2416	Région de la Montérégie

This document provides an overview of the peer group creation process. It presents the 2024 peer group classification and compares the results with previous classifications. Finally, it includes an example illustrating how peer groups can be used to analyze health-related issues.

2 Data

Typically, a set of 23 variables describing the socioeconomic and sociodemographic determinants of health within health regions across Canada is used in the clustering algorithm to generate the peer groups. These variables cover a range of topics including demographic structure, social and economic status, ethnicity, Indigenous status, housing, urbanization, income inequality and labour market conditions. It is important to note that health-related variables were deliberately excluded from the creation of the peer groups.

While some modifications have been made over time, the majority of variables have remained consistent since the creation of the 2000 peer group classification. The 2024 peer group classification uses the same 23 variables as those used in the creation of the 2023 peer groups, with the addition of a new variable— the Gini coefficient. All variables are based on data from the 2021 Census. A detailed list of the variables used in the analysis, along with their respective descriptions, are provided in Table 2.1.

Table 2.1
Variable Definitions

Variable	Description
AVGDWL	Average value of dwelling -owner-occupied, non-farm, non-reserve (Canadian dollars)
EMP	Employment rate (persons aged 25 to 54)
GINICOEFF	Gini index on adjusted household after-tax income
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)

Table 2.1
Variable Definitions

Variable	Description
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)

Note: Variables used in the creation of the 2024 Peer Groups are based on the 2021 Census.

3 Methodology

A non-hierarchical cluster analysis was chosen as the method for forming the peer groups. Cluster analysis, in general, aims to assign observations to groups (or clusters) based on how similar they are to each other, using a distance metric. The objective is to create groups in which the observations are internally similar and distinct from those in other groups — in other words, to form homogeneous and clearly separated clusters. Non-hierarchical clustering algorithms work by dividing a dataset into a predetermined number of non-overlapping groups, using a specific optimization criterion. This approach was considered the most appropriate for meeting the original objectives of the peer group project, which were to apply an empirical method to create a set number of peer groups, each consisting of approximately 5 to 10 health regions.

Traditionally, peer groups were generated in SAS using the **FASTCLUS** procedure. However, for this publication, the groups were created in R using the **cluster** and **flexclust** packages. These packages implement the *k*-means algorithm, which assign observations to a predefined set of *k* clusters. A detailed explanation of *k*-means clustering, and several variants of the method can be found in Johnson and Wicheren (2002). The basic steps of the *k-means* algorithm are:

1. **Initialize:** Select *k* observations to serve as the initial cluster centres (seeds).
2. **Assign:** Allocate each observation to the nearest cluster seed. Once all observations are assigned, update each cluster seed to be the mean of the observations within the cluster. Repeat this step until changes in the cluster centres are minimal or zero.
3. **Finalize:** Assign each observation to its nearest final cluster centre to form the definitive clusters.

3.1 Number of Clusters

One of the main challenges in cluster analysis is determining the appropriate number of initial clusters. Several criteria have been proposed (Everitt et al., 2001) typically involving the optimization of one or more statistical tests. In practice, however, the final decision is often left to the analyst's judgment based on the specific objectives of the study. For the 2024 peer group classification, a maximum of 14 clusters was chosen. This allowed for an average of 7 health regions per peer group,[1] aligning with the study's objectives. The maximum number of clusters used in 2023 was 15.

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 Results

4.1 Standardization of Variables

Variables measured on different scales—or on the same scale but with differing variances—are often standardized to minimize the influence of these disparities. For this exercise, all 24 socioeconomic variables were standardized (mean 0, variance 1) prior to performing the cluster analysis.

Some variables contained missing or zero values, indicating unavailable information for certain health regions. Specifically, the proportion of low-income individuals in private households (LOWPOP) and the proportion of low-income children (LOWKIDS) had missing values, as low-income data are not derived by the Census for the three territories and Indian reserves. These missing values were observed in regions such as "Région des Terres-Cries-de-la-Baie-James" (2418) and the territories. Similarly, the proportion of households spending 30% or more of income on shelter (HOUAFF) and the proportion of owner-occupied dwellings (OWNDWL) also contained missing values for region 2418. All missing values were imputed as zeros before standardizing the variables.

Additionally, the MIZ (Metropolitan Influenced Zone) variable had a value of zero for some health regions. In this case, a zero does not indicate missing data but rather that no large metropolitan area exists within the region. Thus, a zero value for MIZ is valid and was retained in the analysis.

4.2 Creation of Peer groups

To initiate the clustering process, the algorithm was directed to partition the health regions into 14 clusters. However, 6 of the resulting clusters contained fewer than five health regions, suggesting that this number of clusters may be too high for meaningful comparison. The primary objective of forming peer groups is to enable effective comparisons among similar health regions, so the clustering was refined accordingly.

The analysis was rerun with a greater number of iterations, and the K-means algorithm was replaced with K-medians. Unlike K-means, which defines cluster centres by the mean of data points, K-medians uses the median, making it less sensitive to outliers. A final iteration of the clustering was performed with an imposed maximum cluster radius to limit the spread of each group.

The final results of the cluster analysis are presented in Table 4.2.1. This table includes the number of health regions in each peer group and several cluster statistics:

- **Root Mean Square Standard Deviation:** Measures the variability of data points around the cluster centre.
- **Radius:** The maximum Euclidean distance from the cluster centre to any observation within the cluster.
- **Nearest Cluster:** Indicates the most similar peer group based on Euclidean distance.
- **Distance to Nearest Cluster:** Displays the Euclidean distance between the current cluster centre and that of its nearest neighbor.

In this context, each cluster centre is defined by the mean coordinates of all observations within the cluster, and Euclidean distance serves as a standard statistical measure of distance between two points.

Table 4.2.1
Results of final cluster analysis of health regions

Cluster	Frequency	Root Mean Square Standard Deviation	Radius	Nearest Cluster	Distance Between Cluster Centres
A	12	0.52	3.74	I	3.77
B	3	0.38	2.32	C	2.82
C	6	0.34	2.11	B	2.82
D	2	0.53	2.55	H	4.99
E	2	0.47	2.27	F	7.18
F	3	0.74	3.91	E	7.18
G	4	0.33	2.12	J	3.48
H	8	0.48	3.67	K	2.17
I	7	0.68	4.88	A	3.77
J	8	0.40	2.66	G	3.48
K	15	0.45	3.01	H	2.17
L	16	0.43	2.86	N	2.56
M	2	0.49	2.34	I	8.36
N	6	0.38	2.25	L	2.56

Source: Results of the health regions clustering analysis conducted using 24 indicators from the 2021 Census.

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 14 clusters, varying in size from 2 to 16 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 M (Montréal and Toronto). Cluster M was not combined with another cluster since these health regions tend to be very different than other regions across the country.

Cluster B (3 regions) was combined with its nearest neighbor cluster C, producing a cluster of 9 regions. Cluster D (2 regions) was combined with its nearest neighbor cluster H, producing a cluster of 10 regions. Cluster G (4 regions) was combined with its nearest neighbor cluster J, producing a cluster of 12 regions. Finally, clusters E (2 regions) and F (3 regions) were joined together, producing a cluster of 5 regions. The result of collapsing the smaller clusters was that the 14 peer groups produced from the final cluster analysis and presented in Table 4.2.1 were reduced to 10 groups. To maintain continuity in the alphabetical nomenclature of peer groups, the clusters were renamed from A to J. A list of Health regions categorized by the final peer groups can be found in Table 1.1.

4.4 Ontario Health Regions (OHR)

Ontario has two levels of geographic divisions: 6 Ontario Health Regions (OHR) and 34 Public Health Units (PHU). Because of the relationship between these two levels, it was possible to incorporate both into the peer group classification. Information at the PHU level was used to create the peer groups. At the final stage of the cluster analysis, the OHR level geography was incorporated to the existing clusters. The OHR did not affect the placement of the other health regions within the final peer groups. For any analysis involving the peer groups, only one geographic level in Ontario should be used.

Table 4.4.1
Peer Groups for the OHR in Ontario

OHR	Name	Peer Group
3501	West	A
3502	Central	F
3503	Toronto	I
3504	East	A
3505	North East	G
3506	North West	G

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 analyzed using a stepwise discriminant analysis with all 24 variables. A preprocessing step was conducted to detect groups of variables containing redundant information. By automatically removing one variable from each highly correlated pair, we generated a reduced set of predictors in which the remaining variables are largely independent of one another. The **stepclass()** function from the **klaR** package in R was then run in order to sequentially evaluate each candidate predictor's contribution to classification performance by adding or removing variables based on the information criterion (**ability to separate**). A minimum improvement factor was applied to ensure that each variable added to the model provided a meaningful enhancement (≥ 0.05) to its performance. Overall, five variables emerged as the most important predictors. Table 5.1.1 summarizes the results.

Table 5.1.1
Stepwise discriminant analysis of final health region groupings on the 24 variables

Step	Variable	Ability To Separate
1 (Population of Age 0 – 19)	POP20 Added	0.2905
2 (Average Value of Dwelling)	AVGDWL Added	0.4749
3 (Children Living in Low Income Families)	LOWKIDS Added	0.5831
4 (Long-term Unemployment Rate)	LTUNEMP Added	0.6511
5 (5-year Internal Migrants)	MIGMOB Added	0.7085

Source: Results of the discriminant analysis conducted on the 24 indicators from the 2021 Census used for clustering.

5.2 Principal Component Analysis

Principal component analysis (PCA) is a multivariate technique that reduces the number of variables in a dataset to a smaller set of factors called principal components. These components are linear combinations of the original variables and are uncorrelated with each other. They are derived in order of decreasing importance, so that the first few components explain as much of the total variance in the data as possible. Consequently, the first principal component holds the greatest importance, accounting for the largest proportion of total variance in the dataset.

In this study, PCA was performed on 24 socioeconomic variables used in the cluster analysis. The first two principal components accounted for just under 57% of the total variability.

Here is a brief description of the first four components:

1. The first principal component appears to represent factors associated with “urbanicity,” including housing affordability, the proportion of visible minorities, the proportion of immigrants, average dwelling value, and the proportion of the population under the age of 20.
2. The second principal component seems to reflect family profile characteristics, such as the proportion of the population aged 65 and over, the proportion of lone-parent families, the proportion of employed individual aged 25 to 54, the total population in 2021 and the proportion of the Indigenous population.
3. The third principal component can be interpreted as reflecting income inequality, indicated by variables such as the proportion of income received from government transfers, the proportion of low-income children, the proportion of low-income individuals in private households and the unemployment rate.
4. The fourth component is related to the living environment with variables such as population density and owner-occupied private dwellings.

The first six principal components accounted for over 88% of the total variability in the data, demonstrating that 24 variables can be effectively reduced to six factors with minimal loss of information. These results are consistent with the previous peer group classification, indicating that the key variables driving the analysis are remaining fairly stable over time.

5.3 Peer Group Description

The five key variables identified through stepwise discriminant analysis were used to represent each of the clusters. The mean values of these five variables for each peer group are provided in Appendix A. For each variable, several percentiles were calculated and used to classify the peer groups. Values were classified based on the following ranges.

Very High: X > 85th percentile

High: 65th percentile < **X** ≤ 85th percentile

Medium: 35th percentile < **X** ≤ 65th percentile

Low: 15th percentile < **X** ≤ 35th percentile

Very Low: X ≤ 15th 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 I (comprising Montréal and Toronto) is the only group characterized by a very high average value of dwellings, a very high prevalence of children living in low-income families, a very low proportion of 5-year internal migrants, a high long-term unemployment rate and a low proportion of the population aged 0-19.

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

Cluster	Average Value of Dwelling	People Aged 0 - 17 in Low Income Family	Proportion of 5-Year Internal Migrants	Long-term Unemployment Rate	Population Aged 0 - 19 Years
A	High	Very High	Low	Medium	Medium
B	Low	Very Low	Medium	Very Low	Medium
C	Medium	Medium	Medium	Medium	High
D	Medium	Medium	Very Low	Very High	Very High
E	Very Low	Medium	Medium	Very High	Very Low
F	Very High	High	Medium	High	High
G	Medium	Medium	Medium	Low	Medium
H	High	Low	Very High	Medium	Low
I	Very High	Very High	Very Low	High	Low
J	Medium	Medium	Very High	Low	Medium

Source: Summary of the discriminant analysis results conducted on the 24 indicators from the 2021 Census used for clustering.

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

5.4 Geographic Limitation

Each province and territory define 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 both. Considerable variability can exist within health regions regarding health measures due to the lack of geographic homogeneity. This variability should be taken into account when making inferences about a specific region. For example, although health indicators in Vancouver compare favourably with the national averages, this does not imply that residents of Vancouver's downtown core enjoy better-than-average health. This lack of homogeneity in defining health region boundaries complicates the assignment of health regions to peer groups. Such variability can significantly affect how well a specific variable represents the entire region, and in some cases, important defining factors may be overlooked.

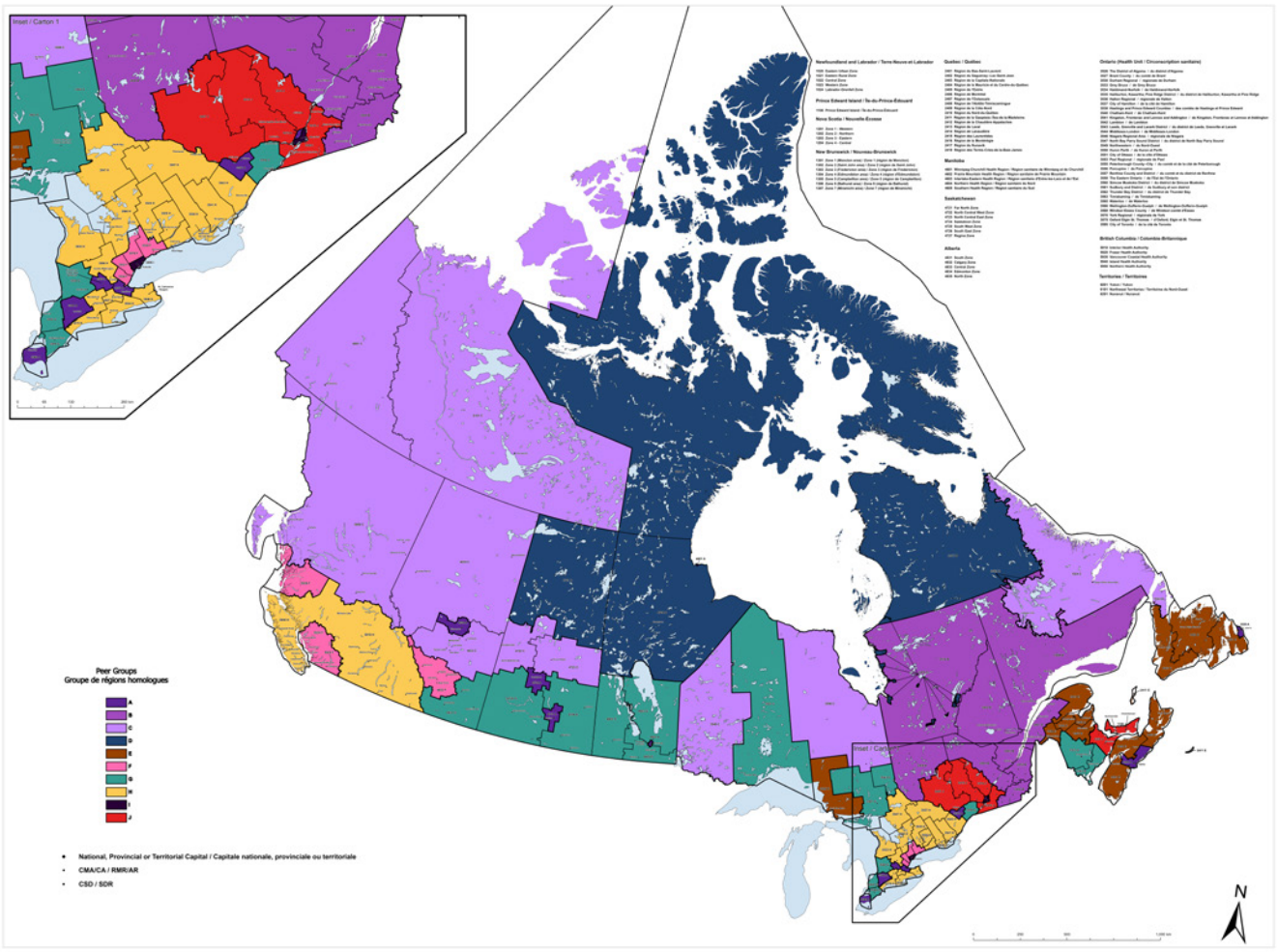
It should also be noted that considerable variability may exist amongst health regions within a peer group regarding the socioeconomic factors used in the cluster analysis. This should be considered when comparing regions within the same peer group. This variability is evident among the 2024 peer groups listed in Appendix A, highlighting 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 10 peer groups. Montréal and Toronto form the smallest cluster due to their significant differences in population size and diversity compared to other health regions, making them unsuitable for inclusion in any other peer group.

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

Map 1
Health Regions and Peer Groups in Canada, 2024



Source: Statistics Canada, Health Regions: Boundaries and Correspondence with Census Geography, catalogue no. 82-402-X.

A larger version of the map is available.

6 Peer Groups in Action

The purpose of this section is to illustrate the usefulness of peer groups. Two valuable but distinct types of analyses are possible using peer groups: comparing health-related indicators between and within peer groups. Since peer groups are formed from regions with similar socioeconomic characteristics, differences between them are expected. Peer groups with more favorable socioeconomic indicators are likely to show better health outcomes. Additionally, estimates from a single peer group can be compared to national averages to assess the overall performance of that group of regions.

A second, and perhaps more relevant, type of analysis involves comparing health regions within the same peer group. Once the effects of socioeconomic factors known to influence health have been accounted for, comparisons based on health status measures become more meaningful.

The example provided in Section 6.1 is a simple illustration of how and when peer groups can be used. The example uses 2024 peer group classification and 2019-2020 Canadian Community Health Survey (CCHS) data. A more in-depth analysis using peer groups is available in the paper "The Health of Canada's Communities" by Margot Shields and Stéphane Tremblay of Statistics Canada (2002).

6.1 Example: Heart Disease

This example examines the prevalence of heart disease among the population aged 18 years and over across different regions of the country. Every respondent in the Canadian Community Health Survey (CCHS) is asked about their heart disease status. The national prevalence of heart disease among adults 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 F is 1.15 percentage points lower than the national average. It is also 1.8 percentage points lower than in Peer Group J. Both differences are statistically significant (p-value<0.01). Peer Group F consists of large cities and suburbs in Ontario, Alberta and British Columbia, characterized by very high population density. This group exhibits a low smoking rate (10.6%), a low heavy drinking rate (15.8%) and an above-average exercise rate (74.2%). Conversely, Peer Group J includes regions with urban and rural areas in Quebec, New Brunswick and Prince Edward Island. This group has a higher smoking rate (16.3%), a higher heavy drinking rate (19.9%) and a lower physical activity rate (66.0%). The differences in these risk factor rates between Peer Groups F and J 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	12	• Mainly urban centres • High average dwelling value • Very high proportion of children living in low income families • Low proportion of 5-year internal migrants	4.60% [4.2%, 5.0%]
B	9	• Regions in Québec outside of Montréal • Low average dwelling value • Very low proportion of children living in low income families • Very low long-term unemployment rate	6.29% [5.7%, 6.9%]
C	10	• Mainly Northern regions in Ontario and British Columbia, rural regions in the Prairies, and Yukon, and Northwest Territories • High proportion of individuals aged 0 to 19 years	5.19% [4.6%, 5.8%]
D	5	• Northern and remote regions with very low population density • Very low proportion of 5-year internal migrants • Very high long-term unemployment rate • Very high proportion of individuals aged 0 to 19 years	3.88% [2.0%, 5.8%]
E	12	• Mainly rural Eastern regions • Very low average dwelling value • Very high long-term unemployment rate • Very low proportion of individuals aged 0 to 19 yearsw	8.32% [7.6%, 9.0%]
F	7	• Large cities and suburbs in Ontario, Alberta and British Columbia • Very high average dwelling value • High proportion of children living in low income families • High long-term unemployment rate • High proportion of individuals aged 0 to 19 years	3.85% [3.5%, 4.2%]
G	15	• Sparsely populated urban-rural mix from coast to coast • Low long-term unemployment rate	6.34% [5.7%, 7.0%]
H	16	• Sparsely populated urban-rural mix in Ontario and British Columbia • High average dwelling value • Low proportion of children living in low income families • Very high proportion of 5-year internal migrants • Low proportion of individuals aged 0 to 19 years	6.06% [5.6%, 6.5%]

Table 6.1.1
Prevalence of Heart Disease by Peer Group

Peer Group	Number of Health Regions	Principal Characteristics	Heart Disease Prevalence
I	2	• Largest metro centres (Toronto, Montreal) • Very high average dwelling value • Very high proportion of children living in low income families • Very low proportion of 5-year internal migrants • High long-term unemployment rate • Low proportion of individuals aged 0 to 19 years	3.97% [3.3%, 4.7%]
J	6	• Regions with urban and rural areas in Québec, New Brunswick and Prince Edward Island • Very high proportion of 5-year internal migrants • Low long-term unemployment rate	5.63% [4.9%, 6.3%]

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 F comprises seven health regions. Table 6.1.2 presents the prevalence of heart disease in each of these regions. All seven regions have a prevalence below the national average of 5.0%. The highest prevalence, at 4.8%, is observed in health region 3570. In contrast, the lowest prevalence of heart disease, at 3.3% and is found in health region 4832.

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

Health Region	Name	Heart Disease Prevalence
3530	Durham Regional Health Unit	4.1% [2.6%, 5.5%]
3536	Halton Regional Health Unit	4.3% [2.7%, 5.9%]
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%]
5920	Fraser Health Authority	3.7% [3.1%, 4.4%]
5930	Vancouver Coastal Health Authority	3.6% [2.7%, 4.6%]

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 prevalence figures presented in the tables 6.1.1 and 6.1.2 can be published without reservation, as they are based on a sufficient number of respondents. The confidence interval can be used to assess the reliability of the estimate itself.

For peer groups that include 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 provincial level to increase the sample size and produce more reliable estimates. In these situations, peer groups offer a useful alternative to provinces.

7 Summary

Due to changes in health region boundaries as of September 2024 and the availability of 2021 Census data, it was necessary to update the 2023 peer group classification. Consistent with the original working paper, the objective was to create a classification that clusters health regions with similar social and economic determinants of health into peer groups. Twenty-four variables covering a broad range of social, economic and demographic factors were used to cluster the health regions.

Starting with an initial set of 14 clusters and ensuring that each cluster contained at least two health regions, the results indicate that six clusters contained fewer than five health regions. Peer groups with fewer than five health regions were combined with their closest neighbour to ensure an adequate number of health regions within a peer group for meaningful comparisons. Cluster I, consisting of Montréal and Toronto, was not merged with another cluster, as these health regions share more similarities among themselves than with others. The final classification comprised 10 peer groups ranging in size from 2 to 16 health regions (excluding Ontario Health Regions (OHR)).

Stepwise discriminant analysis was used to identify the variables that had the greatest influence on the final peer groupings. The five most important variables were the population aged 0 to 19, average dwelling value, children

living in low-income families, long-term unemployment rate, and 5-year internal migrants. Each peer group is characterized by at least one distinctive factor among 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, they allow for more meaningful comparisons between regions. Health indicators can be compared both between and within peer groups. Additionally, peer groups serve as an alternative to provinces when analyses 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.

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
Descriptive statistics for final peer groups

Cluster	Statistics	Average Value of Dwelling	People Aged 0 - 17 in Low Income Family	Proportion of 5-Year Internal Migrants	Long-term Unemployment Rate	Population Aged 0 to 19 Years
A	N	12	12	12	12	12
	MIN	344000	5.7	6.4	8	19.7
	MAX	752000	13.6	19.6	16	25
	Mean	499900	8.22	13.78	11	22.6
	St. Dev	148008.93	2.01	3.36	2	1.66
B	N	9	9	9	9	9
	MIN	170600	1.5	11.9	5	19
	MAX	321600	4.1	21.8	8	25.1
	Mean	232866.67	2.79	16.61	6	21.2
	St. Dev	54401.01	0.86	3	1	1.77
C	N	10	10	10	10	10
	MIN	210400	0	13.6	7	22.1
	MAX	488800	8	21.6	16	28.6
	Mean	320720	4.24	18.82	10	25.2
	St. Dev	85531.4	2.51	2.84	2.43	2.43
D	N	5	5	5	5	5
	MIN	185000	0	8.2	8	37.9
	MAX	470000	8.2	15.3	19	42.9
	Mean	325320	4.26	11.62	14	40.1
	St. Dev	124149.88	4.03	3.1	5	1.95
E	N	12	12	12	12	12
	MIN	135000	1.9	10.8	10	14.7
	MAX	266800	6.8	22.6	21	19.5
	Mean	193050	5.2	16.7	15	17.18
	St. Dev	38650.61	1.35	3.12	4	1.34
F	N	7	7	7	7	7
	MIN	531500	4.3	11.1	8	16.7
	MAX	1548000	10.2	23.8	14	25.6
	Mean	1047071.43	7.09	17.27	11	22.57
	St. Dev	318480.82	1.88	4.45	2	2.88
G	N	15	15	15	15	15
	MIN	222600	4	11.8	6	20.3
	MAX	528500	7.4	24.6	11	30.5
	Mean	323593.33	5.28	19.51	9	22.93
	St. Dev	90810.69	1.11	4.4	2	2.9
H	N	16	16	16	16	16
	MIN	378000	2.7	20.8	8	17
	MAX	802000	6	29.4	16	23.8
	Mean	592737.5	4.14	25.23	11	20.17
	St. Dev	120658.79	0.94	2.9	2	1.99
I	N	2	2	2	2	2
	MIN	638000	11.5	7.6	10	18.6
	MAX	1131000	11.6	10.2	14	20.3
	Mean	884500	11.55	8.9	12	19.45
	St. Dev	348603.64	0.07	1.84	3	1.2
J	N	6	6	6	6	6
	MIN	230000	3.4	18.1	6.3	19.6
	MAX	411200	6.9	31.1	10.3	23.1
	Mean	334200	4.85	24.62	8.1	21.83
	St. Dev	61645.57	1.35	4.43	1.79	1.37

Appendix B
Descriptive summary of final peer groups

Peer Group	Number of Health Regions	Percent of Canadian Population	Principal Characteristics
A	12	18.84%	• Mainly urban centres • High average dwelling value • Very high proportion of children living in low income families • Low proportion of 5-year internal migrants
B	9	7.95%	• Regions in Québec outside of Montréal • Low average dwelling value • Very low proportion of children living in low income families • Very low long-term unemployment rate
C	10	4.51%	• Mainly Northern regions in Ontario and British Columbia, rural regions in the Prairies, and Yukon, and Northwest Territories • High proportion of individuals aged 0 to 19 years
D	5	0.46%	• Northern and remote regions with very low population density • Very low proportion of 5-year internal migrants • Very high long-term unemployment rate • Very high proportion of individuals aged 0 to 19 years
E	12	3.15%	• Mainly rural Eastern regions • Very low average dwelling value • Very high long-term unemployment rate • Very low proportion of individuals aged 0 to 19 years
F	7	23.34%	• Large cities and suburbs in Ontario, Alberta and British Columbia • Very high average dwelling value • High proportion of children living in low income families • High long-term unemployment rate • High proportion of individuals aged 0 to 19 years
G	15	6.62%	• Sparsely populated urban-rural mix from coast to coast • Low long-term unemployment rate
H	16	13.03%	• Sparsely populated urban-rural mix in Ontario and British Columbia • High average dwelling value • Low proportion of children living in low income families • Very high proportion of 5-year internal migrants • Low proportion of individuals aged 0 to 19 years
I	2	12.97%	• Largest metro centres (Toronto, Montreal) • Very high average dwelling value • Very high proportion of children living in low income families • Very low proportion of 5-year internal migrants • High long-term unemployment rate • Low proportion of individuals aged 0 to 19 years
J	6	9.14%	• Regions with urban and rural areas in Québec, New Brunswick and Prince Edward Island • Very high proportion of 5-year internal migrants • Low long-term unemployment rate