

Sources of Data for Rural Health Research: Development of an Inventory of Databases

Health research and planning activities are very much dependent on available data. Secondary data sources provide raw material that can be used for population health research. Composed mainly of routinely collected data for administrative purposes, as well as survey-based data, large-scale, population-level databases can be described as the “essential apparatus” for answering questions geared towards the improvement of the health of a population (Wolfson 1996). As more emphasis on research is placed upon understanding the broader determinants of health rather than simply focussing on medical interventions, secondary data sources are being used more frequently than ever in health research. Given the plans for electronic health records, the systematic collection of an even wider range of data pertaining to the determinants of health may become more feasible. The availability of such databases makes them ideal for analyzing health status and health determinants of Canadians at the population level.

Numerous secondary databases are available for use, but few researchers know how to locate appropriate data sources for their needs. This inventory provides researchers, students and other interested people a tool to support the greater, and more coordinated, use of existing data sources. It will also save them time and effort in identifying useful data sources.

The *Inventory of National Rural Health Research-Related Databases* fits into a broader picture of promoting rural health research. It has the goal of building research capacity within this field. It represents an extensive search for relevant and useful national data sources, and a review of these data series using a common framework. This inventory assesses whether the data are collected, analyzable and released at appropriate geographical levels for rural health research.

This issue of Research in FOCUS on Research is based on the *Inventory of National Rural Health Research-Related Databases*, a component of the “Canada’s Rural Communities: Understanding Rural Health and Its Determinants” study. A 2.5-year study that began in October 2002, the research program is a collaborative undertaking involving the Centre for Rural and Northern Health Research, Health Canada, Statistics Canada, the Canadian Institute for Health Information (CIHI), Dalhousie University, the University of British Columbia, and the Institute for Clinical Evaluative Sciences.

This research is funded by the Canadian Population Health Initiative, which is administered by CIHI.

Building the Inventory

An initial inventory of data sources used for the development of possible indicators of rural health was created in a 1999 study titled *Assessing Rural Health: Towards Developing Health Indicators for Rural Canada* (Pitblado et al. 1999). Building upon this earlier study, a list of existing data sources was compiled. Many data-source websites maintained by Statistics Canada and the Canadian Institute for Health Information (CIHI) were systematically searched with the objective of identifying the types and nature of health and related data available to rural health researchers. Data and documentation from the Data Liberation Initiative of Statistics Canada and other government publications were screened to compile the necessary information for the different data sheets. Finally, knowledgeable individuals from different research centres, health planning agencies, and ministries of health were consulted to improve the inventory data sheets.

some definitions of rural commonly used in Canada and discuss their strengths and limitations.

The functional definitions of rural used in this inventory were taken from *Definitions of "Rural"* (du Plessis et al. 2002). It outlines six alternative ways of defining rural in Canada, five of which are used by the inventory. The building blocks employed by du Plessis et al. have the advantage of assigning individual records in data sources to "neighbourhoods" (i.e., enumeration areas), "communities" (i.e., census subdivisions), or "regions" (i.e., census divisions).

Using postal codes to define rural was not pursued in this inventory due to the many changes to the postal code system since 1996. These changes do not allow researchers to continue to use the second character of the postal code as an indicator of a "rural" area for all provinces or for all parts of some provinces.

Definitions of rural	Building blocks
Census rural areas	Enumeration area (EA)
Rural and small town MIZ	Census subdivision (CSD)
OECD rural communities	Census consolidated subdivision (CCS)
OECD predominantly rural regions	Census division (CD)
Ehrensaft's Beale codes	Census division (CD)

Adapted from du Plessis et al. (2002).

Defining rural

There is no universally accepted definition of rural. For rural health research that is quantitative in nature, the definition depends on the types of geopolitical information collected and the sampling frame and sample size of data sources. Different definitions of rural will generate different populations of rural people. Pong and Pitblado (2001) identify

50 Data Sources Included in the Inventory (as of September 2003)

- Aboriginal Peoples Survey
- Absence From Work Survey
- Annual Demographic Statistics
- Canada Health Survey
- Canada's Alcohol and Other Drugs Survey
- Canadian Cancer Registry
- Canadian Community Health Survey
- Canadian MIS Database
- Canadian Joint Replacement Registry
- Canadian Organ Replacement Register
- Canadian Tobacco Use Monitoring Survey
- Census of Agriculture Custom Retrievals
- Census of Population
- Discharge Abstract Database
- Disease Surveillance Online
- Employment Insurance Statistics
- First Nations and Inuit Regional Health Survey
- Food Expenditure Survey
- Full-time and Part-time Enrolment and Graduates of Postsecondary Programs of Colleges
- General Social Survey
- Health and Activity Limitation Survey
- Health Personnel Database
- Health Promotion Survey
- Health Services Access Survey
- Hospital Mental Health Database

Inventory layout

Each data sheet in this inventory describes the main database characteristics and assesses whether the data from each data source is collected, analyzable and released at appropriate geographical levels for rural health research. It identifies whether the data can be partitioned into rural and non-rural, or into finer geographical categories. Each data sheet uses a common framework that includes such information as:

- ▶ thematic coverage
- ▶ originating agency
- ▶ availability
- ▶ date of data release
- ▶ geographical coverage
- ▶ lowest geographical level for data collection and data release
- ▶ geographical level used in standard reports
- ▶ existing rural/urban flags and how they may be theoretically constructed
- ▶ data elements
- ▶ potential problems in using the data
- ▶ contact information, where available.

Hospital Morbidity Database
Labour Force Survey
Licensed Practical Nurses Database
Maternity Experiences Survey
NATChem/Particulate Matter Database
NATChem/Precipitation Database
National Alcohol and Drug Survey
National Child Care Survey
National Longitudinal Survey of Children and Youth
National Physician Database
National Pollutant Release Inventory
National Population Health Survey (Household Survey)
National Rehabilitation Reporting System
National Survey of Giving, Volunteering and Participating
National Trauma Registry
Participation and Activity Limitation Survey
Registered Nurses Database
Residential Care Facilities Survey
Southam Medical Database
Survey on Aging and Independence
Survey of Smoking Habits
Survey on Smoking in Canada
Traffic Accident Information Database
University Student Information System - Enrolment and Degrees
Vital Statistics
Youth Smoking Survey

Inclusion criteria for data sources considered for this inventory

- ▶ Readily available national data
- ▶ Data relevant to a variety of rural health issues
- ▶ Rural and non-rural geographies
- ▶ Potentially useful data to build population health models

Some Challenges/Opportunities

Several issues became apparent during the compilation of this inventory related to the state of data sources for rural health research in general.

Breadth of data sources

There is a wide range of Canadian databases that can be used by health researchers, but gaps in data sources still exist. For example, there are few data sources in the inventory that deal with features of the physical environment as health determinants. There is also little information about transportation issues, which have been identified as a major problem for access to health care in rural areas.

Undersampling of rural areas

An inherent characteristic of rural Canada is its sparse population. Unfortunately, although some surveys provide reliable estimates at the health region level, the sample size of most was determined with a view to providing reliable estimates at the provincial level. This means that rural figures can often not be extrapolated.

Data suppression

The release of data based on small numbers, which could be of great benefit to rural health researchers, is not permitted due to concerns about statistical reliability, privacy and confidentiality.

Urban/rural flags

Only 15 of the 50 data sources in the inventory include some sort of urban/rural flag. The most frequent definitions of urban/rural used are based on the Statistics Canada “rural and small town” definition or

Census Metropolitan Area (CMA)/Non-CMA codes. Even though these data sources include a rural variable, it is not always possible to use them to conduct rural health research because of data suppression.

Timeliness of data release

Although continuing efforts are being made by data collection agencies to make data available in a timely fashion, such as with the Canadian Community Health Survey, there are other databases where there are significant delays between data collection and release, sometimes up to five years.

Postal codes

The most recent data collections of Statistics Canada and CIHI include postal codes. If respondent confidentiality can be ensured, these codes can be linked to urban and rural areas using postal code conversion files. This would allow for rural health research activities.

References

du Plessis, V et al. (2002). *Definitions of "Rural"*. Agriculture and Rural Working Paper Series, Working Paper No. 61. Ottawa: Statistics Canada, Agricultural Division. www.statcan.ca/english/research/21-601-MIE/21-601-MIE2002061.htm.

Pitblado, JR et al. (1999). *Assessing Rural Health: Toward Developing Health Indicators for Rural Canada*. Sudbury: Centre for Rural and Northern Health Research. www.cranhr.ca/onlrpts.html.

Pong, RW and JR Pitblado (2001). Don't take "geography" for granted! Some methodological issues in measuring geographic distribution of physicians. *Canadian Journal of Rural Medicine* 6(2) : 103-112. collection.nlc-bnc.ca/100/201/300/cdn_medical_association/cjrm/vol-6/issue-2/0103.htm.

Wolfson, MC (1996). *The Role of Large-Scale Population Databases in Health Research*. The Canadian Institute for Advanced Research Program in Population Health, Working Paper No. 60. Toronto.

Accessing the Inventory

The *Inventory of National Rural Health Research-Related Databases* is available in two web-based formats and can be found at the CRaNHR website - www.cranhr.ca under **Resources for Rural Health Research**. A complete PDF version of the inventory can be downloaded from this site, as well as PDF versions of the individual data source sheets. All inventory information is also available in HTML format.

You can help

The compilation of this inventory is intended to be an ongoing project that will be continuously updated as new data and new data sources become available. The current and future editions of this resource will promote the development of common standards for data collection and reporting in Canada. If you notice any information that is incorrect or outdated, or if you are aware of additional information that would help to improve this inventory, please contact:

Denis Heng, Research Associate-Epidemiologist
Phone: 705-675-1151 ext. 4369
Or use the contact information below.

03-A4

Research in **FOCUS** on Research is published by the Centre for Rural and Northern Health Research (CRaNHR), Laurentian University. Each issue is a summary of a study conducted by CRaNHR researchers. As a form of knowledge dissemination and transfer, it is intended to make research accessible to a wider audience.

For further information, please contact:
Centre for Rural and Northern Health Research
Laurentian University
Ramsey Lake Road
Sudbury, Ontario, Canada P3E 2C6

phone: 705-675-1151 ext. 4357
fax: 705-675-4855
e-mail: cranhr@laurentian.ca
URL: www.laurentian.ca/cranhr