Cohort profile: The Canadian Census Health and Environment Cohorts (CanCHECs)

by Michael Tjepkema, Tanya Christidis, Tracey Bushnik and Lauren Pinault

Release date: December 18, 2019
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 STATCAN.infostats-infostats.STATCAN@canada.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

Depository Services Program
- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

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.

An HTML version is also available.

Cette publication est aussi disponible en français.
Cohort profile: The Canadian Census Health and Environment Cohorts (CanCHECs)

by Michael Tjepkema, Tanya Christidis, Tracey Bushnik and Lauren Pinault

Abstract
The Canadian Census Health and Environment Cohorts (CanCHECs) are population-based linked datasets of the household population at the time of census collection. The CanCHECs combine data from respondents to the long-form census or the National Household Survey between 1991 and 2011 with administrative health data (e.g., mortality, cancer incidence, hospitalizations, emergency ambulatory care) and annual mailing address postal codes. The CanCHEC datasets are rich national data resources that can be used to measure and examine health inequalities across socioeconomic and ethnocultural dimensions for different periods and locations. These datasets can also be used to examine the effects of exposure to environmental factors on human health.

Because of their large size, the CanCHECs are an excellent resource for examining rare health outcomes and small population groups. They are ideally suited for environmental health research because of their geographic coverage across all regions of Canada, their long follow-up periods and their linkage to annual postal code history.

Keywords: cohort studies, Canada, mortality, neoplasms, hospitalization, linked health data, census

The reduction and elimination of health inequalities is an ongoing goal of health policy in Canada.1-3 To support and inform progress toward this goal, a sample of 1991 Census respondents and 10 years of mortality data were linked in 2008.4 This dataset was then used to determine the distribution of mortality outcomes across groups defined by income, education, occupation, marital status, language, ethnicity, immigration status, Indigenous identity and disability status. In 2009, approval was granted to add more years of mortality follow-up, and to include cancer incidence data and annual place of residence data.5 The primary purpose of this expanded dataset was to assess the impact of long-term exposure to air pollution on human health, with the objective to inform the development of Canada-wide standards for key criteria pollutants.6 In subsequent years, a 2001 Census linked dataset7 and a 1996 Census linked dataset8 were created and branded as the Canadian Census Health and Environment Cohorts (CanCHECs).

These linked datasets were instrumental in examining health inequalities and the impacts of environmental exposures on mortality. However, they had important differences in population eligibility, linkage methodology, health outcomes and availability of sampling weights. This limited their full analytical potential, particularly their capacity for inter-cohort comparisons. The routine, ongoing measurement of population health status indicators is essential to assess progress in the reduction of inequalities by population and socioeconomic group.9,10 Thus, a more standardized approach to creating the CanCHEC datasets was warranted.

In 2017, a Statistics Canada initiative called the Social Data Linkage Environment (SDLE) was approved. The goals of the SDLE were to promote the innovative use of existing administrative and survey data, to address important questions, and to inform socioeconomic policy through record linkage.11 The SDLE provided an opportunity to create a series of CanCHEC datasets using a common methodology in an efficient and cost-effective manner.

The new CanCHEC datasets (census years 1991 to 2011) provide a rich national data resource that can be used to measure and examine health inequalities (e.g., mortality, cancer incidence, hospitalizations) across socioeconomic and ethnocultural dimensions for different periods and locations. These datasets can also be used to examine the effects of exposure to environmental factors on human health. The main CanCHEC objectives are

● to examine the nature and extent of socioeconomic inequalities in health across Canada over time
● to examine the nature and extent of health inequalities across population groups (e.g., by Indigenous identity, immigration status, ethnicity, language) across Canada over time
● to detect possible elevated risks of mortality and cancer in specific occupations and industries where health risk is known or may be occurring
● to assess the impact of long-term exposure to ambient air pollution on human health
● to examine the impact of other environmental factors (e.g., built environment, social environment, natural environment) on human health

Because of their large size, the CanCHECs are an excellent resource for examining rare health outcomes and small population groups. They are ideally suited for environmental health research because of their geographic coverage across all regions of Canada, their long follow-up periods and their linkage to annual postal code history.
Data resource description

Creation of the cohorts

The CanCHECs are a series of population-based, probabilistically linked datasets that combine data from respondents to the long-form census or the 2011 National Household Survey (NHS) with administrative health data (e.g., mortality, cancer incidence, hospitalizations, emergency ambulatory care) and annual postal code history.

Individuals were eligible to be included in the CanCHECs if they were usual residents of Canada on Census Day (including permanent and non-permanent residents) and if they were in the long-form census12 or 2011 NHS13 records. The institutional population (e.g., those living in nursing homes, penitentiaries, group homes) at the time of census collection was not eligible for CanCHEC inclusion, and the 2011 CanCHEC excludes all collective households (including both institutional and non-institutional). Census data quality reports show that a small proportion of the Canadian population (≤4.3%) is missed in any given census. In general, these individuals are more likely to be young, mobile, low income, homeless or Indigenous.14 Unlike previous census years, participation in the 2011 NHS was voluntary. As a result, the response rate for the NHS was lower (69%) than the response rates for censuses (~95%).

The CanCHEC datasets were created using the SDLE, which facilitates the creation of linked population data files using the Derived Record Depository (DRD). The DRD is a dynamic relational database that contains only basic personal identifiers. It was created from birth, death, immigration and tax files. Eligible census and NHS respondents were probabilistically and deterministically linked to the DRD using standard record linkage methodology. Based on previously linked census–tax datasets,4,7,8 the 1991, 1996 and 2001 CanCHECs were deterministically linked to the DRD using social insurance numbers. The 2006 and 2011 CanCHECs were created within the SDLE from a probabilistic linkage between eligible census and NHS records and the DRD. The linkage rate of in-scope census and NHS records to the DRD differed by age group, marital status, place of residence, socioeconomic status and Indigenous identity.4,7,8 To account for these differences in linkage rates, and to ensure representativeness for each CanCHEC cohort, weights were created from existing census and NHS weights to adjust for non-linkage to the DRD. Bootstrap weights were created to account for variance. Table 1 shows the age eligibility, number of in-scope census and NHS records, cohort size, cohort size compared with the number of in-scope records (percentage), cohort size with annual postal codes, and estimated population for each CanCHEC year. Figure 1 shows the flow of how the CanCHECs were created.

Administrative health data

After the CanCHECs were linked to the DRD, administrative data (previously linked to the DRD) were linked to the CanCHEC datasets. Currently, data from the Canadian Vital Statistics Death15 Database (CVSD), Canadian Cancer Registry16 (CCR), Discharge Abstract Database17 (DAD) and National Ambulatory Care Reporting System18 (NACRS) are available to be linked to the CanCHEC cohorts (the administrative data linked to each CanCHEC varies by province; see Table 2). The linkage rate for applicable years of these health administrative datasets to the DRD ranged from 92.7% to 95.8% for the DAD and NACRS, from 98.1% to 99.1% for the CCR, and from 99.6% to 99.9% for the CVSD.

Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Census Day</td>
<td>June 4</td>
<td>May 14</td>
<td>May 15</td>
<td>May 16</td>
<td>May 10</td>
</tr>
<tr>
<td>Eligible age on Census Day (in years)</td>
<td>25+</td>
<td>19+</td>
<td>19+</td>
<td>0+</td>
<td>0+</td>
</tr>
<tr>
<td>Number of in-scope census/NHS records</td>
<td>3,576,487</td>
<td>4,389,833</td>
<td>4,500,245</td>
<td>6,463,927</td>
<td>6,719,223</td>
</tr>
<tr>
<td>Cohort size (unweighted)</td>
<td>2,643,762</td>
<td>3,564,791</td>
<td>3,536,929</td>
<td>5,871,337</td>
<td>6,499,183</td>
</tr>
<tr>
<td>Cohort size compared with number of in-scope records (percentage)</td>
<td>73.9</td>
<td>81.2</td>
<td>78.6</td>
<td>90.8</td>
<td>96.7</td>
</tr>
<tr>
<td>Cohort size (unweighted) with postal code history†</td>
<td>2,642,110</td>
<td>3,564,771</td>
<td>3,536,293</td>
<td>5,873,695</td>
<td>6,499,183</td>
</tr>
</tbody>
</table>

† Postal code file is available only for those aged 15 or older.

NHS = National Household Survey

Notes: Estimated population size is randomly rounded to a base of five.


Table 2

<table>
<thead>
<tr>
<th>Health outcome data, follow-up years and counts by Canadian Census Health and Environment Cohort (CanCHEC) version</th>
<th>CanCHEC</th>
<th>CVSD (N=)</th>
<th>CCR (N=)</th>
<th>DAD (N=)</th>
<th>NACRS (N=)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(734,734)</td>
<td>(503,301)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(725,399)</td>
<td>(571,655)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001 (version 3)</td>
<td>2001 to 2016</td>
<td>1992 to 2015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(532,252)</td>
<td>(493,325)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(436,407)</td>
<td>(491,083)</td>
<td>(8,923,516)</td>
<td>(35,974,416)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(210,089)</td>
<td>(403,271)</td>
<td>(8,416,871)</td>
<td>(35,834,597)</td>
<td></td>
</tr>
</tbody>
</table>


Annual postal codes

The T1 Personal Master File (T1PMF) was the principal data source for addresses—including postal codes—in the DRD. Every resident of Canada who earns taxable income is required to complete an income tax return, or T1 form, after the end of the year in which the income was received. Therefore, the T1PMF includes almost all individuals who filed an individual T1 tax return for the reference year. However, some late filers may not be included. A filer’s address may be updated later in the year if they communicated with the Canada Revenue Agency. The DRD also obtains addresses from other files, such as the Canadian Vital Statistics Birth Database, the CVSD, the Canada Child Tax Benefit Identifier file and the Immigrant Landing File.

Annual postal codes (available as early as 1981) were linked to the cohort members aged 15 or older within the DRD. While the earlier CanCHECs had at least one annual postal code linked to each record, this was not the case for the more recent CanCHECs for two reasons. First, the 2006 and 2011 CanCHECs include long-form census respondents younger than age 15. Second, only those who agreed on both the 2006 Census and 2011 NHS to have their information linked to their tax records could have annual postal codes linked to their records.

Between 1981 and 2016, the overall percentage of cohort members aged 20 or older with at least one annual postal code was between 93% and 95%, depending on the CanCHEC year.

CanCHEC characteristics

Each CanCHEC consists of data from long-form census records or NHS records, mortality records, and cancer incidence records—with hospitalization records and emergency ambulatory care records available for some provinces, starting with the 2006 CanCHEC. The long-form censuses and the NHS collected a wide range of information on topics such as education, labour, income, language, immigration, ethnocultural diversity, Indigenous identity and housing (Table 3). Administrative health outcomes linked to the CanCHECs include

---

**Table 3**

**Summary of key analytical variables by data source**

<table>
<thead>
<tr>
<th>Data source</th>
<th>Key variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Census (long form) and NHS</td>
<td>Family composition and marital status, education, labour, occupation, income and source of income, Indigenous identity, ethnic diversity and immigration, country of birth, housing, language, collective dwellings</td>
</tr>
<tr>
<td>CVSD</td>
<td>Age at time of death, date of death, underlying cause of death</td>
</tr>
<tr>
<td>CCR</td>
<td>Age at time of cancer diagnosis, primary cancer diagnosis, characteristics of cancer</td>
</tr>
<tr>
<td>DAD</td>
<td>Inpatient hospitalizations, age, admission and discharge dates, length of stay, procedures and diagnosis</td>
</tr>
<tr>
<td>NACRS</td>
<td>Emergency ambulatory care services, age at admission, date of arrival, emergency room diagnosis</td>
</tr>
</tbody>
</table>

NHS = National Household Survey  
CVSD = Canadian Vital Statistics Death Database  
CCR = Canadian Cancer Registry  
DAD = Discharge Abstract Database  
NACRS = National Ambulatory Care Reporting System

Sources: Statistics Canada, census (long form), NHS, CVSD 2015 and CCR 2015; Canadian Institute for Health Information, DAD and NACRS.
underlying and contributing causes of death, date of death, primary cancer diagnoses, date of cancer diagnoses, inpatient hospitalizations, admission and discharge dates, length of stay, procedures and diagnoses, emergency room diagnoses, and date of emergency room arrival.

Because exclusion criteria were applied during the creation of the CanCHECs (i.e., CanCHEC records include only the non-institutional population at baseline who has been enumerated by the census or NHS), the cohort population is healthier than the Canadian population. Linkage error (estimated to be small) is also present (e.g., false negative or missed links to a health outcome file) and could result in underestimations (e.g., underestimated mortality rates, hospitalization rates, cancer incidence rates). When compared with official mortality tables, which are based on the total Canadian population, CanCHEC mortality rates are similar for younger age groups and lower for the oldest age groups—particularly among respondents aged 85 and older (Table 4). This pattern is consistent with the age distribution of the institutional population. The proportion of the population living in institutions—and thus excluded from the CanCHECs—is lowest in the younger age groups (less than 1%) and highest in the oldest age groups (greater than 5% for those aged 80 and older) (data not shown).

### Table 4

<table>
<thead>
<tr>
<th>Mortality rate per 1,000, and absolute mortality rate differences by sex, age group and data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1 to 4 years</td>
</tr>
<tr>
<td>5 to 9 years</td>
</tr>
<tr>
<td>10 to 14 years</td>
</tr>
<tr>
<td>15 to 19 years</td>
</tr>
<tr>
<td>20 to 24 years</td>
</tr>
<tr>
<td>25 to 29 years</td>
</tr>
<tr>
<td>30 to 34 years</td>
</tr>
<tr>
<td>35 to 39 years</td>
</tr>
<tr>
<td>40 to 44 years</td>
</tr>
<tr>
<td>45 to 49 years</td>
</tr>
<tr>
<td>50 to 54 years</td>
</tr>
<tr>
<td>55 to 59 years</td>
</tr>
<tr>
<td>60 to 64 years</td>
</tr>
<tr>
<td>65 to 69 years</td>
</tr>
<tr>
<td>70 to 74 years</td>
</tr>
<tr>
<td>75 to 79 years</td>
</tr>
<tr>
<td>80 to 84 years</td>
</tr>
<tr>
<td>85 to 89 years</td>
</tr>
<tr>
<td>90 years and older</td>
</tr>
</tbody>
</table>

### Note:
- CanCHEC rates are based on five-year mortality follow-up periods; Canadian population rates are based on five-year averages (1991 to 1995; 1996 to 2000; 2001 to 2005; 2006 to 2010; 2011 to 2016).
Annual postal code history is an important feature of the CanCHECs, particularly for environmental health research and attaching environmental exposures. More than two-thirds of CanCHEC cohort members aged 20 or older had complete postal code histories, with the 2011 CanCHEC having the highest completion rate, at 74% (Table 5). The percentage of CanCHEC members who moved (i.e., changed postal codes) in any given year was between 6.0% and 10.0%, depending on the cohort. The percentage who moved to a different region of Canada, based on the first letter of the postal code, was between 1.1% and 1.7% per year.

Data resource use

The CanCHECs have been widely used over the past decade across a range of topics, resulting in more than 65 peer-reviewed journal articles since 2008. Research conducted using CanCHEC data can be grouped into four broad research themes: health outcomes by socioeconomic position such as income, education and occupational skill level; health outcomes by specific occupation, occupational group and job characteristic; health outcomes by population group, such as Indigenous peoples and immigrants; and the measurement of the effect of environmental exposures on health.

Socioeconomic position

CanCHEC data have been used extensively to examine socioeconomic indicators and their relationship with life expectancy, cause-specific mortality, cancer incidence and hospitalizations. Research findings highlighting these inequalities have been cited in published reports by the Chief Public Health Officer of Canada2 and the Organisation for Economic Co-operation and Development.19 In general, higher levels of educational attainment, household income and occupational skill are all linked to increased longevity,20 lower mortality rates for most causes of death,21-24 lower rates of cancer incidence for many cancers25,26 and lower rates of opioid-related hospitalizations.27 Further, mortality inequalities by socioeconomic status appear to have increased between 1991 and 2016 for both all-cause mortality28 and health-adjusted life expectancy,29 suggesting that the reduction of health inequalities remains a relevant issue in Canada. In general, the CanCHEC datasets have been identified as an important resource to examine the effects of socioeconomic status on mortality.30 Going forward, the CanCHECs are well suited to study the intersection of multiple dimensions of socioeconomic status and population groups across different periods.

Occupation

Research identifying elevated risks of mortality and cancer in specific occupations has also benefitted from the CanCHECs. For example, cancer incidence rates by cancer site have been estimated for police officers, firefighters, armed forces personnel,31 welders,32 agricultural workers,33 workers employed in more sedentary occupations34 and workers who may be exposed to whole-body vibration.35 For mortality outcomes, suicide by major and minor occupation group has been estimated.36 All-cause and cardiovascular mortality rates have been estimated by job overqualification status,37 and avoidable mortality rates have been estimated by occupation.38 These findings can contribute to occupational surveillance programs and can help identify at-risk occupations.

Population groups

The Truth and Reconciliation Commission of Canada identified enhanced data as a priority in their Calls to Action.39 CanCHEC data can respond to this priority by overcoming previous methodological challenges in estimating health outcomes such as life expectancy and cancer burden among First Nations people, Métis and Inuit living in Canada. CanCHEC research has shown that life expectancy is substantially and consistently shorter for First Nations people, Métis and Inuit compared with the non-Indigenous population across all CanCHEC years (1991 to 2011).40 The magnitude of mortality inequalities between non-Indigenous and First Nations people and Métis varies considerably by cause of death.41 Using the 1991 CanCHEC and the linked cancer data, research has shown that First Nations people and Métis have lower cancer survival rates and higher incidence rates for some cancer sites.42-44 Going forward, the CanCHEC data can be an excellent resource for ongoing surveillance of mortality, cancer and hospitalization outcomes for Indigenous people in Canada. These results will be able to identify inequalities not previously estimated and measure any progress made to reduce these gaps.

Researchers have used CanCHEC data to examine mortality and cancer outcomes for immigrants. Several studies have demonstrated a healthy immigrant effect—lower mortality and cancer rates among immigrants compared with the Canadian-born population—with the greatest advantage observed among recent immigrants.45-48 These same studies have shown significant variation across immigrant subgroups by region.
of birth, time in country, cause of death and cancer site. In the future, it will be possible to examine health outcomes by immigrant category over time as new data sources, such as the Immigrant Landing File, are added to the DRD.

Environmental health

During the past two decades, the CanCHEC datasets have been used as a basis for influential work in Canadian environmental health, particularly in the field of air pollution. The availability of annual postal code history has enabled researchers to follow cohort members through time with a reasonably accurate residential history, and attach them to various environmental exposures. Researchers have examined the effects of exposure to long-term ambient air pollution on cause-specific mortality. This is of particular importance internationally, since deleterious effects have been observed even at the lower levels of exposure that are typical in Canada. In addition to providing valuable input for reaching international scientific consensus on air pollution and health, these studies have helped national (i.e., the Public Health Agency of Canada) and international bodies identify future challenges for reducing air pollution.

CanCHEC data have been used to link ambient ultraviolet radiation to increasing rates of melanoma, a finding that has been informative for public health outreach campaigns in Canada. In addition to environmental hazards, CanCHEC data have been used in environmental benefit studies. Living near greenspace or water has been linked to reductions in mortality. Findings from these studies have been included in a knowledge translation tool to share environmental health assessments across Canada.

Strengths and weaknesses

The CanCHEC datasets have many strengths. By their nature, the datasets fill a gap by linking individual-level national administrative health data that lack socioeconomic and ethnocultural identifiers with individual-level census long-form data that contain these variables. As a result, administrative health outcomes can be examined across characteristics such as income; education; occupation; language; ethnicity; First Nations, Métis and Inuit status; and immigration status. The large sample size and extended follow-up periods allow for the examination of health outcomes for small populations (e.g., First Nations people, Métis, Inuit) and for rare or uncommon disease outcomes, especially those related to occupational exposures.

The CanCHECs offer a breadth of health data—from emergency ambulatory care visits to death records—that can be used to answer a variety of research questions. The DRD is regularly updated to include new years of administrative data and new administrative data sources. Likewise, the CanCHECs can be updated continuously to prolong follow-up periods and to include new data sources.

Trends over time both within and between CanCHECs can now be examined because of the consistent SDLE-based linkage methodology used to create the cohorts. The cohorts can be examined in concert, or stacked for a combined analysis, which is useful for examining rare health outcomes (e.g., rare causes of death).

Annual historical postal codes from 1981 onward provide an indication of where cohort members live year after year, and more importantly, can be used to attach environmental exposures to a person’s place of residence. Postal codes provide a reasonably accurate approximation of residence location in most cities, particularly for urban street addresses or apartment buildings, where they can be geocoded within about 200 metres. However, postal codes in rural areas and those that represent businesses or post office boxes can be quite far from the residence’s location (on average, five to six kilometres), creating a limitation to this geocoding approach.

Researchers should consider the following aspects of the CanCHECs when assessing whether the data are appropriate for their needs. Census characteristics are collected only at the time of census collection, and may not reflect changes in a person’s socioeconomic characteristics over their lifespan. No information regarding baseline health (with the exception of activity limitation) or health behaviour risk factors—such as smoking and physical activity—is collected by the census. However, population-level and individual-level indirect adjustment techniques have been applied using Canadian Community Health Survey data. Moreover, administrative health data are available for some years prior to Census Day for certain CanCHECs. Specifically, the CCR has data starting in 1992, the DAD in 2000 and the NACRS in 2002.

The CanCHEC datasets grossly underestimate infant mortality (deaths that occur within the first year after birth) because about three-quarters of all infant deaths occur within the first 28 days. This makes census enumeration uncertain. As a result, life expectancy at birth cannot be reliably estimated using the CanCHECs.

The institutional population on Census Day and individuals not enumerated by the census are excluded from the CanCHECs. Because of these exclusions, the population represented by CanCHEC data should be considered a healthier population than the general population. With sampling weights, the CanCHEC data can be considered representative of the non-institutional population at the time of census collection, but bias might exist if those missing from the cohort differed systematically from those who were included.

The annual postal code file is primarily derived from mailing addresses provided on T1 tax records, and may not be associated with a person’s place of residence. For example, the mailing address may represent a post office box, an accountant or lawyer’s office, a parent’s address for young adults, or a child’s address for elderly parents. Moreover, missing postal codes are imputed whenever possible. For example, if a person does not file taxes for two years, but uses the same mailing address in their tax filings in the years before and after this gap, the SDLE
will automatically impute this address as the mailing address for the missing years. This assumes that the person did not move to a different address in the intervening years before returning to the initial address.

The introduction of the voluntary 2011 NHS was an important methodological change from previous long-form censuses. It is not fully known how this change affects the comparability with CanCHEC cycles that are based on mandatory long-form censuses. However, the development of weights may mitigate some of the comparability issues among census cycles.

Finally, for analysis across CanCHEC years, it is important to consider changes over time to the census questionnaire, and to be aware that some census variables may not be available in a given year, or may have been modified over time.

**Data resource access**

Approved researchers can access the CanCHECs in research data centres (RDCs) and in the Federal Research Data Centre. Information on the RDC Program, including the application process and guidelines, is available at www.statcan.gc.ca/eng/rdc/index.

**Acknowledgements**

The authors are grateful to all the people, too numerous to name individually, who have worked on these census linkages over the past two decades. The authors would like to acknowledge Russell Wilkins (Statistics Canada, retired), who was instrumental in developing the original 1991-to-2001 Census Mortality Follow-Up Study, and Richard Burnett (Health Canada), who envisioned the enhancement and use of these data for ambient air pollution research. The authors would also like to thank the Canadian Institute for Health Information and Health Canada for their financial assistance in creating some of the earlier linkages.

**References**


