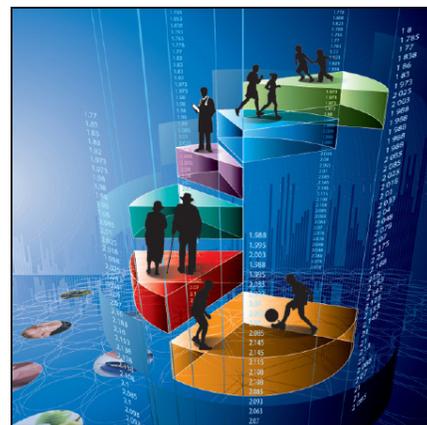


Health Reports

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



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 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.

Published by authority of the Minister responsible for Statistics Canada

© Her Majesty the Queen in Right of Canada as represented by the Minister of Industry, 2019

All rights reserved. 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.

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

DOI: <https://www.doi.org/10.25318/82-003-x201901200003-eng>

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 ($\leq 4.3\%$) 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
Description of Canadian Census Health and Environment Cohorts (CanCHECs)

	1991	1996	2001	2006	2011
Census Day	June 4	May 14	May 15	May 16	May 10
Eligible age on Census Day (in years)	25+	19+	19+	0+	0+
Number of in-scope census/NHS records	3,576,487	4,389,833	4,500,245	6,463,927	6,719,223
Cohort size (unweighted)	2,643,762	3,564,791	3,536,929	5,871,337	6,499,183
Cohort size compared with number of in-scope records (percentage)	73.9	81.2	78.6	90.8	96.7
Cohort size (unweighted) with postal code history†	2,642,110	3,564,771	3,536,210	3,783,695	3,769,270
Estimated population (weighted cohort)	17,450,900	21,036,890	22,247,440	31,232,960	32,852,265

† 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.

Sources: 1991, 1996 and 2001 CanCHEC (version 3), 2006 and 2011 CanCHEC (version 1).

Table 2
Health outcome data, follow-up years and counts by Canadian Census Health and Environment Cohort (CanCHEC) version

CanCHEC	CVSD (N=)	CCR (N=)	DAD (N=)	NACRS (N=)
1991 (version 3)	1991 to 2016 (734,734)	1992 to 2015 (503,301)
1996 (version 3)	1996 to 2016 (725,399)	1992 to 2015 (571,655)
2001 (version 3)	2001 to 2016 (532,252)	1992 to 2015 (493,325)
2006 (version 1)	2006 to 2016 (436,407)	1992 to 2015 (491,083)	2000/2001 to 2016/2017 (8,923,516)	2002/2003 to 2017/2018 (35,974,416)
2011 (version 1)	2011 to 2016 (210,089)	1992 to 2015 (403,271)	2000/2001 to 2016/2017 (8,418,871)	2002/2003 to 2017/2018 (35,834,597)

.. not available for a specific reference period

CVSD = Canadian Vital Statistics Death Database

CCR = Canadian Cancer Registry (excludes Quebec data from 2011 onward)

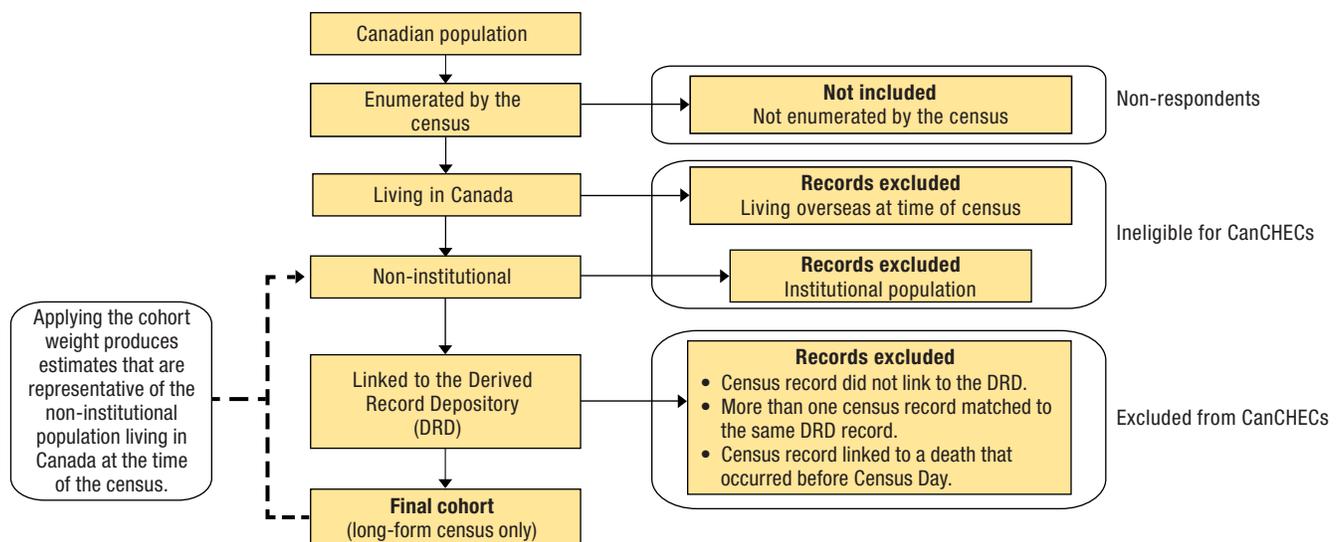
DAD = Discharge Abstract Database (excludes Quebec data and Nunavut data for 2002/2003)

NACRS = National Ambulatory Care Reporting System

Notes: NACRS includes data for Prince Edward Island starting in 2003/2004, Nova Scotia starting in 2003/2004, Ontario starting in 2002/2003, Manitoba starting in 2009/2010, Saskatchewan starting in 2010/2011 to 2017/2018, Alberta starting in 2010/2011, British Columbia starting in 2002/2003 excluding 2010/2011, and Yukon starting in 2004/2005.

Sources: 1991, 1996 and 2001 CanCHEC (version 3), 2006 and 2011 CanCHEC (version 1).

Figure 1
Derivation of the Canadian Census Health and Environment Cohorts (CanCHECs)



Source: Canadian Census Health and Environment Cohorts.

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

Table 3
Summary of key analytical variables by data source

Data source	Key variables
Census (long form) and NHS	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
CVSD	Age at time of death, date of death, underlying cause of death
CCR	Age at time of cancer diagnosis, primary cancer diagnosis, characteristics of cancer
DAD	Inpatient hospitalizations, age, admission and discharge dates, length of stay, procedures and diagnosis
NACRS	Emergency ambulatory care services, age at admission, date of arrival, emergency room diagnosis

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.

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

Cohort profile: The Canadian Census Health and Environment Cohorts (CanCHECs) • Statistics Canada Data Resource Profile

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
Mortality rate per 1,000, and absolute mortality rate differences by sex, age group and data source

	CanCHEC population					Canadian population					Mortality rate difference (CanCHEC minus Canadian population)				
	1991	1996	2001	2006	2011	1991	1996	2001	2006	2011	1991	1996	2001	2006	2011
Males															
1 to 4 years	0.2	0.2	0.4	0.3	0.2	0.2	0.2	0.0	0.0
5 to 9 years	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.0	0.0
10 to 14 years	0.2	0.1	0.2	0.2	0.2	0.1	0.1	0.0	0.0
15 to 19 years	0.5	0.4	0.8	0.7	0.6	0.5	0.4	0.0	0.0
20 to 24 years	...	0.7	0.7	0.7	0.7	1.1	1.0	0.8	0.8	0.7	...	-0.2	-0.1	0.0	0.0
25 to 29 years	0.9	0.8	0.7	0.8	0.7	1.2	0.9	0.8	0.8	0.8	-0.2	-0.1	-0.1	0.0	-0.1
30 to 34 years	1.1	0.9	0.8	0.8	0.6	1.4	1.1	0.9	0.8	0.8	-0.2	-0.2	-0.2	-0.1	-0.2
35 to 39 years	1.5	1.1	0.9	0.9	0.9	1.7	1.4	1.2	1.1	1.0	-0.2	-0.3	-0.3	-0.2	-0.1
40 to 44 years	1.9	1.6	1.4	1.4	1.2	2.2	2.0	1.8	1.6	1.5	-0.3	-0.4	-0.4	-0.2	-0.3
45 to 49 years	2.9	2.5	2.3	2.2	2.0	3.3	2.9	2.8	2.5	2.3	-0.3	-0.4	-0.5	-0.3	-0.3
50 to 54 years	5.0	4.0	3.9	3.6	3.1	5.3	4.6	4.3	4.1	3.6	-0.3	-0.6	-0.5	-0.5	-0.5
55 to 59 years	8.4	7.1	6.1	5.6	5.4	9.0	7.8	6.9	6.4	5.9	-0.5	-0.8	-0.9	-0.8	-0.5
60 to 64 years	14.3	12.1	10.4	8.9	8.2	14.9	13.0	11.5	10.0	9.2	-0.6	-0.9	-1.1	-1.1	-1.0
65 to 69 years	23.6	20.2	16.7	14.5	13.0	24.2	21.7	18.4	15.8	14.0	-0.5	-1.5	-1.6	-1.4	-1.0
70 to 74 years	36.6	33.0	27.3	23.5	20.8	38.1	35.3	29.9	25.5	22.7	-1.5	-2.3	-2.6	-2.1	-2.0
75 to 79 years	57.0	51.4	44.8	38.3	33.7	61.5	56.8	49.2	42.3	37.1	-4.5	-5.4	-4.4	-4.0	-3.4
80 to 84 years	88.2	85.0	72.4	64.9	55.7	97.6	94.1	81.2	71.5	64.5	-9.5	-9.1	-8.8	-6.6	-8.8
85 to 89 years	132.7	129.8	118.2	106.8	95.3	154.2	150.7	136.9	119.4	112.3	-21.5	-20.9	-18.7	-12.6	-17.0
90 years and older	197.0	210.1	201.2	192.4	174.4	236.8	251.4	232.6	216.8	214.0	-39.8	-41.3	-31.4	-24.4	-39.6
Females															
1 to 4 years	0.1	0.1	0.3	0.2	0.2	0.2	0.2	-0.1	-0.1
5 to 9 years	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.0	0.0
10 to 14 years	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.0	0.0
15 to 19 years	0.2	0.2	0.3	0.3	0.3	0.3	0.2	-0.1	0.0
20 to 24 years	...	0.3	0.3	0.3	0.3	0.4	0.3	0.3	0.3	0.3	...	0.0	0.0	0.0	0.0
25 to 29 years	0.4	0.3	0.3	0.3	0.3	0.4	0.4	0.3	0.3	0.3	0.0	-0.1	0.0	0.0	0.0
30 to 34 years	0.5	0.4	0.5	0.4	0.4	0.5	0.5	0.5	0.4	0.4	-0.1	-0.1	0.0	0.0	0.0
35 to 39 years	0.7	0.6	0.7	0.6	0.5	0.8	0.7	0.7	0.6	0.6	-0.1	-0.1	-0.1	-0.1	-0.1
40 to 44 years	1.1	1.0	0.9	0.9	0.8	1.2	1.2	1.1	1.0	0.9	-0.1	-0.2	-0.2	-0.1	-0.1
45 to 49 years	1.9	1.6	1.5	1.6	1.3	2.0	1.8	1.8	1.7	1.5	-0.1	-0.2	-0.3	-0.1	-0.3
50 to 54 years	2.9	2.6	2.4	2.4	2.2	3.2	2.9	2.7	2.6	2.4	-0.3	-0.3	-0.3	-0.3	-0.3
55 to 59 years	4.7	4.2	3.9	3.7	3.6	5.2	4.7	4.3	4.0	3.8	-0.5	-0.5	-0.3	-0.3	-0.3
60 to 64 years	7.5	6.6	6.3	5.7	5.4	8.1	7.5	6.9	6.2	5.8	-0.6	-0.9	-0.6	-0.5	-0.4
65 to 69 years	11.6	10.5	9.8	9.4	8.4	12.9	12.0	11.0	9.9	9.0	-1.3	-1.5	-1.2	-0.5	-0.6
70 to 74 years	18.3	16.9	16.0	14.6	13.1	20.5	19.6	17.9	16.2	14.7	-2.2	-2.7	-1.9	-1.6	-1.6
75 to 79 years	28.8	27.2	25.3	23.6	20.7	34.8	33.1	30.3	27.1	24.4	-6.0	-5.9	-5.0	-3.5	-3.7
80 to 84 years	47.2	43.9	40.9	39.7	35.4	59.6	58.3	52.8	47.8	44.3	-12.4	-14.4	-11.9	-8.1	-8.8
85 to 89 years	76.6	74.9	69.6	68.6	59.5	103.9	103.4	97.0	86.7	80.9	-27.3	-28.5	-27.4	-18.1	-21.3
90 years and older	137.5	139.9	132.1	140.1	125.6	200.2	209.2	196.8	182.8	178.4	-62.7	-69.2	-64.7	-42.7	-52.8

... not applicable

CanCHEC = Canadian Census Health and Environment Cohort

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).

Sources: 1991, 1996 and 2001 CanCHEC (version 3), 2006 and 2011 CanCHEC (version 1); deaths and mortality rates, by age group (Table 13-10-0710-01).

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 cancers^{25,26} 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

Table 5
Percentage of follow-up years with a postal code for persons aged 20 or older, living in Canada, by CanCHEC year

Percentage complete	CanCHEC year				
	1991	1996	2001	2006	2011
100% complete	68.6	66.5	69.4	69.8	74.1
At least 80% complete	87.6	86.4	89.4	88.9	90.5
At least 60% complete	94.1	93.8	95.9	95.6	96.2
At least 40% complete	96.8	97.2	98.5	98.3	98.3
At least 20% complete	98.3	98.8	99.4	99.4	99.1

Notes: Denominator is based on cohort members in the postal code file who were living in Canada and were aged 20 or older for each follow-up year.

CanCHEC = Canadian Census Health and Environment Cohort.

Sources: 1991, 1996 and 2001 CanCHEC (version 3), 2006 and 2011 CanCHEC (version 1).

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.⁴⁹ Researchers have examined the effects of exposure to long-term ambient air pollution on cause-specific mortality.⁵⁰⁻⁵³ 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,^{54,55} these studies have helped national (i.e., the Public Health Agency of Canada)⁵⁶ and international⁵⁷ bodies identify future challenges for reducing air pollution.

CanCHEC data have been used to link ambient ultraviolet radiation to increasing rates of melanoma,⁵⁸ a finding that has been informative for public health outreach campaigns in Canada.⁵⁹

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.^{60,61} Findings from these studies have been included in a knowledge translation tool to share environmental health assessments across Canada.⁶²

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.⁶³ 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 char-

acteristics 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.^{64,65} 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.⁶⁶ 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.⁴⁹ 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 centres⁶⁷ (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

1. Health Canada. [2003 First Ministers' Accord on Health Care Renewal: A Progress Report](#). Available at: <https://www.canada.ca/en/health-canada/services/health-care-system/health-care-system-delivery/federal-provincial-territorial-collaboration/first-ministers-meeting-year-plan-2004/2003-accord-health-care-renewal-progress.html>. Accessed July 30, 2019.
2. Health Canada. *Chief Public Health Officer's Report on the State of Public Health in Canada, 2008: Addressing Health Inequalities* (Health Canada, Catalogue HP2-10/2008) Ottawa: Minister of Health, 2008.
3. Public Health Agency of Canada and the Pan-Canadian Public Health Network. *Key Health Inequalities in Canada: A National Portrait* (Public Health Agency of Canada, Catalogue HP35-109/2018E-PDF) Ottawa: Public Health Agency of Canada, 2018.
4. Wilkins R, Tjepkema M, Mustard C, Choinière R. The Canadian census mortality follow-up study, 1991 through 2001. *Health Reports* 2008; 19(3): 25-43.
5. Peters PA, Tjepkema M, Wilkins R, et al. Data resource profile: 1991 Canadian Census Cohort. *International Journal of Epidemiology* 2013; 42(5): 1319-26. doi: 10.1093/ije/dyt147.
6. Statistics Canada. [Air pollution study: Linkage of 1991 Census of Population, Canadian Mortality Database and Canadian Cancer Database follow-up study](#). Approved microdata linkages, 2009 submission. Available at: <https://www.statcan.gc.ca/eng/record/2009#a23>. Accessed July 30, 2019.
7. Pinault L, Finès P, Labrecque-Synnott F, et al. The 2001 Canadian Census-Tax-Mortality Cohort: a 10-year follow-up. *Analytical Studies: Methods and References* No. 003 (Catalogue 11-633-X) Ottawa: Statistics Canada, 2016.
8. Christidis T, Labrecque-Synnott F, Pinault L, et al. The 1996 CanCHEC: Canadian Census Health and Environment Cohort profile. *Analytical Studies: Methods and References* No. 013 (Catalogue 11-633-X) Ottawa: Statistics Canada, 2018.
9. Global Burden of Disease 2017 Mortality Collaborators. Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2018; 392(10159): 1684-1735.
10. Public Health Agency of Canada, Pan-Canadian Public Health Network, Statistics Canada, and Canadian Institute of Health Information. [Pan-Canadian Health Inequalities Data Tool, 2017 Edition](#). Available at: <https://health-infobase.canada.ca/health-inequalities/data-tool>. Accessed July 30, 2019.
11. Statistics Canada. [Social Data Linkage Environment \(SDLE\)](#). Available at: <https://www.statcan.gc.ca/eng/sdle/index>. Accessed July 30, 2019.
12. Statistics Canada. [Census of Population](#). Available at: <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3901>. Accessed July 30, 2019.
13. Statistics Canada. [National Household Survey](#). Available at: <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5178>. Accessed July 30, 2019.
14. Statistics Canada. *Census Technical Report: Coverage Census of Population, 2011* (Catalogue 98-303-X) Ottawa: Statistics Canada, 2015.
15. Statistics Canada. [Vital Statistics – Death Database \(CVSD\): Detailed Information for 2015](#). Available at: <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3233>. Accessed July 30, 2019.
16. Statistics Canada. [Canadian Cancer Registry \(CCR\): Detailed Information for 2015](#). Available at: <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3207>. Accessed July 30, 2019.
17. Canadian Institute for Health Information. [Discharge Abstract Database Metadata \(DAD\)](#). Available at: <https://www.cihi.ca/en/discharge-abstract-database-metadata>. Accessed July 30, 2019.
18. Canadian Institute for Health Information. [National Ambulatory Care Reporting System \(NACRS\) Metadata](#). Available at: <https://www.cihi.ca/en/national-ambulatory-care-reporting-system-metadata>. Accessed July 30, 2019.
19. Murtin F, Mackenbach J, Jasilionis D, d'Ercole MM. Inequalities in longevity by education in OECD countries. *OECD Statistics Working Papers* 2017/02. Available at: <http://dx.doi.org/10.1787/6b64d9cf-en>. Accessed July 30, 2019.
20. Tjepkema M, Wilkins R. Remaining life expectancy at age 25 and probability of survival to age 75 by socio-economic status and Aboriginal ancestry. *Health Reports* 2011; 22(4): 31-36.

Cohort profile: The Canadian Census Health and Environment Cohorts (CanCHECs) • Statistics Canada Data Resource Profile

21. Tjepkema M, Wilkins R, Long A. Cause-specific mortality by education in Canada: a 16-year follow-up study. *Health Reports* 2012; 23(3): 23-31.
22. Tjepkema M, Wilkins R, Long A. Cause-specific mortality by income adequacy in Canada: a 16-year follow-up study. *Health Reports* 2013; 24(7): 14-22.
23. Tjepkema M, Wilkins R, Long A. Cause-specific mortality by occupational skill level in Canada: a 16-year follow-up study. *Chronic Diseases and Injuries in Canada* 2013; 33(4): 195-203.
24. Tjepkema M, Wilkins R, Long A. Socio-economic inequalities in cause-specific mortality: a 16-year follow-up study. *Canadian Journal of Public Health* 2013; 104(7): e472-478.
25. Mitra D, Shaw A, Tjepkema M, Peters P. Social determinants of lung cancer incidence in Canada: a 13-year prospective study. *Health Reports* 2015; 26(6): 12-20.
26. Bushnik T, Evans WK. Sociodemographic characteristics associated with thyroid cancer risk in Canada. *Health Reports* 2018; 29(10): 3-11.
27. Carrière G, Garner R, Sanmartin C. Social and economic characteristics of those experiencing hospitalizations due to opioid poisonings. *Health Reports* 2018; 29(10): 23-28.
28. Marshall-Catlin E, Bushnik T, Tjepkema M. Trends in mortality inequalities among the adult household population between 1991 and 2011. *Health Reports* 2019; 30(12): 11-17.
29. Bushnik T, Tjepkema M, Martel L. Socioeconomic disparities in life and health expectancy among the household population in Canada. *Health Reports* 2020 (in press).
30. Bourbeau R, Ouellette N. Trends, patterns, and differentials in Canadian mortality over nearly a century (1921-2011). *Canadian Studies in Population* 2016; 43(1-2): 48-77.
31. Harris MA, Kirkham TL, MacLeod JS, et al. Surveillance of cancer risks for firefighters, police, and armed forces among men in a Canadian census cohort. *American Journal of Industrial Medicine* 2018; 61(10): 815-823. doi: 10.1002/ajim.22891.
32. Macleod JS, Harris MA, Tjepkema M, et al. Cancer risks among welders and occasional welders in a national population-based cohort study: Canadian Census Health and Environment Cohort (CanCHEC). *Safety and Health at Work* 2017; 8(3): 258-266. doi: 10.1016/j.shaw.2016.12.001.
33. Kachuri L, Harris MA, MacLeod JS, et al. Cancer risks in a population-based study of 70,570 agricultural workers: results from the Canadian Census Health and Environment Cohort (CanCHEC). *BMC Cancer* 2017; 17(1): 343. doi: 10.1186/s12885-017-3346-x.
34. Pahwa M, Harris MA, MacLeod J, et al. Sedentary work and the risks of colon and rectal cancer by anatomical sub-site in the Canadian Census Health and Environment Cohort (CanCHEC). *Cancer Epidemiology* 2017; 49: 144-151. doi: 10.1016/j.canep.2017.06.004.
35. Jones MK, Harris MA, Peters PA, et al. Prostate cancer and occupational exposure to whole-body vibration in a national population-based cohort study. *American Journal of Industrial Medicine*. 2014; 57(8): 896-905. doi: 10.1002/ajim.22354.
36. Mustard CA, Bielecky A, Etches J, et al. Suicide mortality by occupation in Canada, 1991-2001. *Canadian Journal of Psychiatry* 2010; 55(6): 369-76.
37. Smith BT, Smith PM, Etches J, Mustard CA. Overqualification and risk of all-cause and cardiovascular mortality: evidence from the Canadian census mortality follow-up study (1991-2001). *Canadian Journal of Public Health* 2012; 103(4): e297-302.
38. Mustard CA, Bielecky A, Etches J, et al. Avoidable mortality for causes amenable to medical care, by occupation in Canada, 1991-2001. *Canadian Journal of Public Health* 2010b; 101(6): 500-6.
39. Truth and Reconciliation Commission of Canada. *Honouring the Truth, Reconciling for the Future: Summary of the Final Report of the Truth and Reconciliation Commission of Canada*. Winnipeg: Truth and Reconciliation Commission of Canada, 2015.
40. Tjepkema M, Bushnik T, Bougie E. Life expectancy of the Indigenous household population in Canada. *Health Reports* 2019; 30(12): 3-10.
41. Tjepkema M, Wilkins R, Senécal S, Guimond É. Mortality of Métis and Registered Indian adults in Canada: an 11-year follow-up study. *Health Reports* 2009; 20(4): 31-51.
42. Withrow DR, Pole JD, Nishri ED, et al. Cancer survival disparities between First Nations and non-Aboriginal adults in Canada: follow-up of the 1991 Census Mortality Cohort. *Cancer Epidemiology, Biomarkers & Prevention* 2016. doi: 10.1158/1055-9965.EPI-16-0706.
43. Mazereeuw MV, Withrow DR, Diane Nishri E, et al. Cancer incidence among First Nations adults in Canada: follow-up of the 1991 Census Mortality Cohort (1992-2009). *Canadian Journal of Public Health* 2018; 109(5-6): 700-709. doi: 10.17269/s41997-018-0091-0.
44. Mazereeuw MV, Withrow DR, Nishri ED, et al. Cancer incidence and survival among Métis adults in Canada: results from the Canadian census follow-up cohort (1992-2009). *CMAJ* 2018; 190(11): e320-6. doi: 10.1503/cmaj.170272.
45. Omariba DW, Ng E, Vissandjée B. Differences between immigrants at various durations of residence and host population in all-cause mortality, Canada 1991-2006. *Population Studies* 2014; 68(3): 339-57.
46. Omariba DW. Immigration, ethnicity, and avoidable mortality in Canada, 1991-2006. *Ethnicity and Health* 2015; 20(4): 409-436.
47. McDonald JT, Franworth M, Liu Z. Cancer and the healthy immigrant effect: a statistical analysis of cancer diagnosis using a linked census-cancer registry administrative database. *BMC Public Health* 2017; 17(1): 296. doi: 10.1186/s12889-017-4190-2.
48. Ng E. The healthy immigrant effect and mortality rates. *Health Reports* 2011; 22(4): 25-29.
49. Bérard-Chagnon J. Comparison of place of residence between the T1 Family File and the census: evaluation using record linkage. *Demographic Documents* No. 13 (Catalogue 91F0015M) Ottawa: Statistics Canada, 2017.
50. Crouse DL, Peters PA, van Donkelaar A, et al. Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study. *Environmental Health Perspectives* 2012; 120(5): 708-14. doi: 10.1289/ehp.1104049.
51. Crouse DL, Peters PA, Hystad P, et al. Ambient PM_{2.5}, O₃, and NO₂ exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environmental Health Perspectives* 2015; 123(11): 1180-6. doi: 10.1289/ehp.1409276.
52. Pinault LL, Weichenthal S, Crouse DL, et al. Associations between fine particulate matter and mortality in the 2001 Canadian Census Health and Environment Cohort. *Environmental Research* 2017; 159: 406-15. doi: 10.1016/j.envres.2017.08.037.
53. Burnett R, Chen H, Szyszkwicz M, et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences of the United States of America* 2018; 115(38): 9592-7. doi: 10.1073/pnas.1803222115.
54. Thurston GD, Kipen H, Annesi-Maesano I, et al. 2016. A joint ERS/ATS policy statement: what constitutes an adverse health effect of air pollution? An analytical framework. *European Respiratory Journal* 2017; 49: 1600419.
55. Global Burden of Disease 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* 2016; 388: 1659-724.
56. Public Health Agency of Canada. *Chief Public Health Officer's Report on the State of Public Health in Canada 2017: Designing Healthy Living* Ottawa: Public Health Agency of Canada, 2017.
57. Greenbaum, DS. The clean air act: substantial success and the challenges ahead. *Annals of the American Thoracic Society* 2017; 15(3): 296-7.

58. Pinault L, Bushnik T, Fioletov V, et al. The risk of melanoma associated with ambient summer ultraviolet radiation. *Health Reports* 2017; 28(5): 3-11.
59. Tenkate T. Slip, slop, slap and wrap: turning up the heat on sun safety and skin cancer prevention. Presentation available from *Public Health Ontario*, 2018.
60. Crouse DL, Pinault L, Balram A, et al. Urban greenness and mortality in Canada's largest cities: a national cohort study. *Lancet Planet Health* 2017; 1(7): e289-97. doi: 10.1016/S2542-5196(17)30118-3. Erratum in: *Lancet Planet Health*. 2017; 1(8): e315.
61. Crouse DL, Balram A, Hystad P, et al. Associations between living near water and risk of mortality among urban Canadians. *Environmental Health Perspectives* 2018; 126(7): 077008. doi: 10.1289/EHP3397.
62. de Souza, RJ, Arora R, Gauvin L, et al. Development of an on-line interactive map to display environmental health assessments of Canadian communities: knowledge-translation to support collaborations for health. *Cities & Health* 2018; 2(2): 123-9.
63. Khan S, Pinault L, Tjepkema M, Wilkins R. Positional accuracy of geocoding from residential postal codes versus full street addresses. *Health Reports* 2018; 29(2): 3-9.
64. Shin HH, Cakmak S, Brion O, et al. Indirect adjustment for multiple missing variables applicable to environmental epidemiology. *Environmental Research* 2014; 134: 482-7.
65. Erickson AC, Brauer M, Christidis T, et al. Evaluation of a method to indirectly adjust for unmeasured covariates in the association between fine particulate matter and mortality. *Environmental Research* 2019; 175: 108-16. doi: 10.1016.
66. Statistics Canada. [Table 13-10-0712-01 Infant Deaths and Mortality Rates, by Sex](https://www150.statcan.gc.ca/t1/tb1/en/tv.action?pid=1310071201). Available at: <https://www150.statcan.gc.ca/t1/tb1/en/tv.action?pid=1310071201>.
67. Statistics Canada. [Research Data Centre Program](https://www.statcan.gc.ca/eng/rdc/index). Available at: <https://www.statcan.gc.ca/eng/rdc/index>. Accessed July 30, 2019.