



## Data quality, concepts and methodology

<b>1.</b>	<b>Regional Health Indicators.....</b>	<b>2</b>
1.1	Health Region Level Population Estimates.....	3
<b>2.</b>	<b>Health Status Indicators Based on Vital Statistics (Statistics Canada).....</b>	<b>3</b>
2.1	Provincial Vital and Cancer Statistics.....	3
2.2	Regional-Level Vital Statistics Indicators.....	4
2.2.1	Regional Level Data Quality Measures: Confidence Intervals .....	4
2.2.2	Age-Standardized Rates .....	5
2.2.3	Geographic Coding (Geo-Coding) to Health Regions .....	6
2.2.4	Birth Statistics .....	6
2.2.5	Life Expectancy .....	6
2.2.6	Disability-Free Life Expectancy .....	7
2.2.7	Disability-Adjusted Life Expectancy (DALE).....	8
2.2.8	Deaths Due to Medically Treatable Diseases.....	9
2.2.9	Potential Years of Life Lost .....	9
<b>3.</b>	<b>Indicators Based on Cancer Incidence (Statistics Canada).....</b>	<b>11</b>
3.1	Cancer Incidence.....	11
<b>4.</b>	<b>Indicators Based on Statistics Canada Surveys.....</b>	<b>12</b>
4.1	National Population Health Survey .....	12
4.2	National Population Health Survey -- Northern Component.....	12
4.3	Canadian Community Health Survey .....	14
4.4	National Longitudinal Survey of Children and Youth.....	14
4.5	Bootstrapping .....	15
<b>5.</b>	<b>Indicators Based on Crime Data (Statistics Canada) .....</b>	<b>15</b>
<b>6.</b>	<b>Indicators Based on Labour Force Data (Statistics Canada).....</b>	<b>16</b>
<b>7.</b>	<b>Indicators Based on Census Data (Statistics Canada) .....</b>	<b>17</b>
<b>8.</b>	<b>Health System Indicators (Canadian Institute for Health Information - CIHI).....</b>	<b>18</b>
8.1	Hospitalization Data and Rates (CIHI).....	18
8.2	Physician Data (CIHI) .....	20
8.3	National Health Expenditure Database (CIHI).....	20
<b>Appendix .....</b>		<b>22</b>
Population estimates .....		22
Statistics Canada Methodology.....		22
Provincial Methodologies .....		24



## **Data Quality, Concepts and Methodology**

### **1. Regional Health Indicators**

The methodology used for these indicators was designed to maximize inter-regional and inter-provincial comparability given the characteristics of available national datasets. For this reason, there may be differences between definitions, data sources, and extraction procedures used in local, regional, or provincial/territorial reports when compared to those described here. In addition, discrepancies may exist due to on-going updates to databases.

Rates are standardized wherever possible to facilitate comparability across provinces/regions and over time.

Indicators in this release that are based on hospitalization records and produced by the Canadian Institute for Health Information (CIHI) are limited to health regions with population greater than 75,000.

For most data sources (with the exception of Census and Demographic population estimates), health region level data are not available for some northern health regions in Manitoba and Saskatchewan which have small populations. To avoid suppression in these areas where small numbers or sample size impact on data quality, data have been grouped with neighbouring regions, as follows:

1. In Manitoba, Burntwood Regional Health Authority (4680) is combined with Churchill Regional Health Authority (4690) and referred to as 'Burntwood/Churchill (4685).'
2. In Saskatchewan, Mamawetan Churchill River Regional Health Authority (4711) is combined with Keewatin Yatthé Regional Health Authority (4712) and Athabasca Health Authority (4713) and referred to as 'Athabasca/Keewatin/Mamawetan (4714).

## 1.1 Health Region Level Population Estimates

Population estimates for health regions were produced by Statistics Canada (Demography Division) for all provinces, except Alberta and British Columbia. Alberta population estimates from Alberta Health and Wellness and British Columbia population estimates were provided by BC Stats. See Appendix for methodology.

## 2. Health Status Indicators Based on Vital Statistics (Statistics Canada)

### 2.1 Provincial Vital and Cancer Statistics

Health indicators based on vital and cancer statistics are produced at the Canada, province and territorial level only, with long time series. These indicators may have different methodologies compared to the regional health indicators (refer to section 2.2). Data on provincial health and on regional health may be the same indicator, but the numbers or rates may differ because of their methodologies. One key difference is that the provincial indicators are based on single years of data, whereas regional level data are based on multiple year averages (refer to section 2.2). For this reason, in addition to certain additional methodological differences, comparisons between these two sources are not recommended.

These provincial health indicators include the Canada/province/territory-only time series data for Life Expectancy, Low Birth Weight, Age-standardized Mortality Rates, Infant Mortality, Potential Years of Life Lost and Cancer Incidence.

Age-standardized mortality and cancer incidence rates were based on place of residence. The formula for age-standardization is presented in a later section entitled "Age-standardized mortality rates". Cancer incidence data from 2001 to 2004 are estimates produced by Health Canada.

Life expectancy is calculated using the Greville method, a widely recognized method of constructing a life table<sup>1</sup>. These provincial/territorial life expectancy data were based on single years of mortality and population and were abridged life tables (i.e., 5 year age-sex groupings). Although their methodologies differ, the Greville, Chiang and Keyfitz methods of calculating life expectancy yield similar results<sup>2</sup>. There are no special notes for the provincial vital statistics indicators of low birth weight and infant mortality outside of what is described in the Definitions and Data Sources document.

Potential years of life lost (PYLL) was calculated in the same fashion as the regional-level indicators of the same name, as described in section 2.2.9.

---

1. Greville TNE. Short methods of constructing abridged life tables. *The Record of American Institute of Actuaries* 1943; 32(65):29-42, Part 1

2. Ng Edward and Gentleman Jane F, "The Impact of Estimation Method and Population Adjustment on Canadian Life Table Estimates", *Health Reports* 1995, Vol. 7, No.3, pp.15-22

## 2.2 Regional-Level Vital Statistics Indicators

- ◆ Rates are based on place of residence for indicators derived from birth and death events.
- ◆ Indicators presented in this product (with the exception of province-only indicators, described above) which were derived from vital statistics, are based on three years of data in both numerator and denominator. For low birth weight, three years of birth data are used in both the numerator and denominator. For infant and perinatal mortality, three years of death or stillbirth data are divided by the same three years of birth data. For mortality, three years of death data (e.g., 2000 to 2002) are divided by three times the mid-year (e.g., 2001) population estimate. In all vital statistics table titles, the year mentioned simply refers to the middle year.

### 2.2.1 Regional Level Data Quality Measures: Confidence Intervals

- ◆ All data presented have an associated 95% confidence interval (CI). The confidence interval illustrates the degree of variability associated with a rate. Wide confidence intervals indicate high variability, thus, these rates should be interpreted and compared with due caution. Some age-standardized rates were suppressed due to both a very small underlying count plus extremely high variability. Confidence intervals can also be used to determine whether a rate in one health region is statistically below, above or no different than the rate for the same indicator in another health region.
- ◆ The confidence intervals for the age-standardized rates were produced using the variance derived using the Spiegelman method.<sup>3</sup>

$$Variance = \sum_x \left( \frac{P_x^s}{P^s} \right)^2 \cdot \frac{m_x(1-m_x)}{P_x}$$

where  $P^s$  is the standard population (refer to section 2.2.2),  $P_x^s$  is the age-specific standard population,  $x$  is the age group (using 5-year age groups),  $P_x$  is the population estimate for the corresponding age group,  $m_x$  is the mean age-specific crude mortality rate and  $n$  is the number of years of data used.

Note that when using  $n$  years of data,  $m_x = \frac{\sum_{i=1}^n d_{xi}}{nP_x}$ , where  $d_{xi}$  is the number of deaths in age group  $x$  in year  $i$ .

- ◆ The confidence intervals for the crude count, crude rate and birth-related data were produced via the Fleiss method.<sup>4</sup> Take note that the lower confidence interval (CI) in this formula is constrained by zero, which means the difference between the rate and the lower CI is not always equal to the difference between the rate and the upper CI.

---

3. Spiegelman M. Introduction to Demography, Revised Edition. Cambridge Massachusetts: Harvard University Press, 1968. p 113, Formula (4.29).

4. Fleiss JL, Statistical Methods for Rates and Proportions, 2nd Ed, Wiley and Sons, NY, 1981. pg. 14, Formula (1.26) and (1.27).

$$P_L = \frac{(2np + c^2 - 1) - c\sqrt{c^2 - (2 + (1/n)) + 4p(nq + 1)}}{2(n + c^2)} \quad \text{for the lower limit and}$$

$$P_U = \frac{(2np + c^2 + 1) + c\sqrt{c^2 + (2 - (1/n)) + 4p(nq - 1)}}{2(n + c^2)} \quad \text{for the upper limit}$$

where  $n$  is the number of events,  $p$  is the proportion or rate,  $c$  is the standard error (1.96 at 95% confidence) and  $q = (1 - p)$ . Remember that  $n$  is comprised of  $x$  years worth of data, and  $p = n/pop$ , where  $pop$  is  $x$  years worth of life-years.

## 2.2.2 Age-Standardized Rates

- ◆ Mortality rates, with the exception of crude rates, potential years of life lost (PYLL) and infant and perinatal mortality, as well as cancer incidence and certain CIHI-based data, are age-standardized using the direct method, and the 1991 Canadian Census population structure. The use of a standard population results in more meaningful rate comparisons, as it adjusts for variations in population age distributions over time and across different geographic areas.

Age (in years)	Standard Population	Age (in years)	Standard Population
<1	403,061	45-49	1,674,153
1-4	1,550,285	50-54	1,339,902
5-9	1,953,045	55-59	1,238,441
10-14	1,913,115	60-64	1,190,217
15-19	1,926,090	65-69	1,084,588
20-24	2,109,452	70-74	834,024
25-29	2,529,239	75-79	622,221
30-34	2,598,289	80-84	382,303
35-39	2,344,872	85-89	192,410
40-44	2,138,891	90+	95,467

Source: Statistics Canada Cat. No. 84F0208XPB, Causes of Death 1997, Appendix 3

The formula for age-standardized death rate  $r$  is:

$$r = \sum_{i=1}^{20} w_i \cdot \left( \frac{d_i}{p_i} \right)$$

Where for age group  $i$ ,  $d_i$  is the age-sex specific death count,  $p_i$  is the population size for a given cause of death and geographical area, and  $w_i$  is the weight for that group. Note that the same weight is used for each sex. To yield a rate per 100,000 population,  $r$  is multiplied by 100,000.

### 2.2.3 Geographic Coding (Geo-Coding) to Health Regions

- ◆ Birth and death data have been linked to health regions using postal codes reported with **place of residence** and converted to census geography using the automated geo-coding system (PCCF Plus<sup>5</sup>) developed by the Health Statistics Division of Statistics Canada. These data were then aggregated to health region based on correspondence files<sup>6</sup> developed by the Health Statistics Division with the cooperation of provincial Ministries of Health, Alberta Treasury and BC Stats.
- ◆ Where postal codes were not available or invalid, additional steps were taken to assign records to health regions using the census subdivision codes for place of residence recorded on the national birth and death database. Stillbirth data, used to calculate perinatal mortality, were linked to health regions solely using census subdivision codes.

### 2.2.4 Birth Statistics

Birth data on the Vital Statistics Database for Ontario are underestimated due to incomplete files. Thus, birth-related indicators (low birth weight, infant mortality and perinatal mortality), particularly for Ontario, should be interpreted with caution.

### 2.2.5 Life Expectancy

This variable was calculated using the Chiang methodology for abridged life tables. The estimates are based on three years (e.g., 2000-2002) of mortality data and the mid-year population estimates, as described above. Abridged life tables use five-year age groupings of both population and mortality rate inputs (as opposed to single year age breakdown). Since there is more variability in the number of events by age in smaller geographic areas, abridged life tables are more suitable for the adaptation to a sub-provincial level (health region). Chiang's method in particular was chosen because it was relatively easy to adapt to the health region level data and included the calculation of standard error (in this case, addressing the variability of deaths from one year to the next).

---

5. *Postal Code Conversion File –Plus*, 82F0086XDB

6. *Health regions: boundaries and correspondence with census geography*, 82-42-XIE  
Health Indicators, December 2005

## 2.2.6 Disability-Free Life Expectancy

Estimates of disability-free life expectancy are calculated using Sullivan's method<sup>7</sup>. Essentially, the latter generalizes Chiang's method<sup>8</sup>.

Sullivan's method is based on activity limitation rates within a population, according to sex and age group, in the calculation of life expectancy with disability. In the case of people living in health institutions, it was assumed that everyone had at least one activity limitation. For people living in other types of institutions, the hypothesis established is that the activity limitation rate by age group and sex was identical to the population in private households.

Disability-free life expectancy represents the difference between life expectancy and life expectancy with disability. The standard deviations of disability-free life expectancy estimates (and consequently the upper and lower limits of the confidence intervals associated with these estimates) are based on Colin Mathers' method<sup>9</sup>. This method takes into account both the stochastic fluctuations in the observed death rates and the sampling variability of the activity limitation rates.

*Note:* The disability data for DFLE came from the 1996 Census of Population. Questions on disability in the Census of Population are generally used to capture the sample of post-censal Health and Activity Limitations Survey. Because of the decision not to conduct this survey in 1996, data on disability from the Census of population of 1996 were neither verified nor imputed. More precisely, no validation was undertaken to check the completeness or consistency of the data, and as a result, no corrections to the data were made. In addition, the data were not adjusted to account for population undercounts.

DFLE estimates will vary according to both the concepts from which they are based and the surveys from which the data are extracted.

**DFLE (Volume 2001, No's. 1 and 2):** For these issues, disability was defined as "having any activity limitation or handicap".

**DFLE (Volume 2001, No. 3 and beyond):** Disability is defined as "having an activity limitation that affects activities at home, work or at school". This differs from previous Health Indicators issues by excluding limitations that only affect activities other than home, work or school as well as respondents who stated that they had some form of handicap other than an activity limitation.

---

7. Sullivan, DF. A single index of mortality and morbidity. *HSMHA Health Reports* 86 (April 1971) : 347-354

8. Chiang, CL. *The Life Table and its Applications*. Robert E. Krieger Publishing Co., Malabar, Florida, 1984: 316

9. Mather, C. Health Expectancies in Australia 1981 and 1988. Australian Government Publishing Service, Canberra, 1991: 117

## 2.2.7 Disability-Adjusted Life Expectancy (DALE)

Disability-adjusted life expectancy (DALE) is similar to DFLE, in that they are both measures of quality of life lived and both are based on mortality and disability data. However, DALE is an expectation of life weighted to account for four health states defined in terms of disability. These health states are, in order of greatest to least weight: (1) having no activity limitations; (2) having activity limitations in leisure time activities and/or transportation; (3) having activity limitations at work, home and/or school; and (4) institutionalization in a health care facility. Specifically, state #1 has a weight of 1.0; state #2 has a weight of 0.8; state #3 has a weight of 0.65; and state #4 has a weight of 0.5. The sum of life expectancies of persons in a specific age group within a given geography, based on their health states, produces the value of DALE for that specific age group.

The calculation of the confidence intervals for DALE is based on Colin Mathers' method. Specifically, for any particular age group,

$$\sigma_{DALE} = \sqrt{(1.0)^2 \cdot (\sigma_{LE_{state1}})^2 + (0.8)^2 \cdot (\sigma_{LE_{state2}})^2 + (0.65)^2 \cdot (\sigma_{LE_{state3}})^2 + (0.5)^2 \cdot (\sigma_{LE_{state4}})^2}$$

Where  $\sigma$  is the standard error, LE denotes life expectancy and 'state  $n$ ' refers to the specific health state.

## 2.2.8 Deaths Due to Medically Treatable Diseases

- ◆ The definitions of medically treatable diseases were taken from a paper written by WW Holland<sup>10</sup>. This was based on earlier work from JRH Charlton<sup>11</sup>. The types of medically treatable diseases mentioned in Charlton originally came from a paper by DD Rutstein<sup>12</sup>.
- ◆ All results were age-standardized according to the age group considered for reasonable odds of survival. These age-standardized rates per 100,000 reflect these age groups, not the total population.
- ◆ The method of calculating confidence intervals was the Spiegelman method (refer to section 2.2.1).

## 2.2.9 Potential Years of Life Lost

- ◆ In this publication, death was considered premature if the person died before age 75. This is more reflective of life expectancies in recent years and is more reflective of international standards. Many previous Statistics Canada publications provide PYLL data based on death before age 70. Additionally, PYLL can be presented as an age-standardized rate or as a crude rate; in this publication, it is presented as a crude rate. As well, the denominator can be based on population aged 0 to 74 or for the total population. In this publication, the denominator is based on the former.
- ◆ In this publication, a PYLL rate was produced, where the weights are taken as proportions of the years lost per death within each age group over the total years lost in all age groups. Each death event is multiplied by its age-specific weight. The sum of all these values represents the total PYLL. The PYLL rate is PYLL per 100,000 population aged 0 to 74. The use of weights allows for the calculation of confidence intervals. The confidence intervals for each PYLL rate were produced by the Spiegelman method (refer to section 2.2.1).

---

10. Holland WW and EC Working Group on Health Services and Avoidable Deaths (1997), "European Community Atlas of Avoidable Death 1985 to 1989", Oxford, Oxford Medical Publications, Commission of the European Communities Health Services Research Series, no.9, p. 371

11. Charlton JRH, "Avoidable deaths and diseases as monitors of health promotion", pp. 467-479, in Measurement in health promotion and protection, Copenhagen and Albany NY: World Health Organization and the International Epidemiological Association, 1987

12. Rutstein DD, "Monitoring progress and failure: sentinel health events (unnecessary diseases, disabilities and untimely deaths)", pp. 195-212, in Measurement in health promotion and protection, Copenhagen and Albany NY: World Health Organization and the International Epidemiological Association, 1987

AGE GROUP	YEARS LOST	WEIGHT
Under 1	74.9	74.9/636.9
1-4	72.0	72.0/636.9
5-9	67.5	67.5/636.9
10-14	62.5	62.5/636.9
15-19	57.5	57.5/636.9
20-24	52.5	52.5/636.9
25-29	47.5	47.5/636.9
30-34	42.5	42.5/636.9
35-39	37.5	37.5/636.9
40-44	32.5	32.5/636.9
45-49	27.5	27.5/636.9
50-54	22.5	22.5/636.9
55-59	17.5	17.5/636.9
60-64	12.5	12.5/636.9
65-69	7.5	7.5/636.9
70-74	2.5	2.5/636.9
<b>SUM</b>	<b>636.9</b>	<b>1.0</b>

This publication only presents PYLL rates based on the sum of all age groups. Thus, the rate is calculated as follows:

$$Rate = \frac{((\sum PYLL) \cdot w)}{3 \cdot POP}$$

Where  $\sum PYLL$  is the sum of PYLL for ages 0 to 74 for the three years of data,  $w$  is a weight of 1, and 'POP' is the population aged 0-74 for the middle year of the three years.

To calculate the age-specific PYLL rates:  $Rate_i = \frac{\sum (PYLL_i \cdot w_i)}{3 \cdot POP_i}$

where  $i$  is the specific 5-year age group.

### 3. Indicators Based on Cancer Incidence (Statistics Canada)

Latest health region level rates are based on the boundaries in effect as of June 2005.

#### 3.1 Cancer Incidence

The Canadian Cancer Registry (CCR) is a central database located at Statistics Canada that contains patient-oriented information about diagnosis of cancers in Canada. Data on the incidence of cancer are collected by the provincial and territorial cancer registries. The information is used for descriptive and analytic epidemiological studies: to identify risk factors for the cancer; to plan, monitor and evaluate a broad range of cancer control programs (e.g., screening); and for health services and economic research and planning.

- ◆ Cancer incidence is based on place of residence at time of diagnosis.
- ◆ Rates contained in this publication have been tabulated using the July 2005 tabulation file, the World Health Organization, International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3) and the International Agency for Research on Cancer (IARC) rules for determining multiple primaries sites.
- ◆ Cancer incidence data in this product are based on three years of data (e.g., 1999 to 2001) averaged over the population estimate of the middle year (e.g., 2000). Table titles associated with these data reflect the mid-point of the three-year period being averaged (i.e., 2000).
- ◆ All data presented have an associated 95% confidence interval (CI). The confidence interval illustrates the degree of variability associated with a rate. Wide confidence intervals indicate high variability, thus, these rates should be interpreted and compared with due caution. Some age-standardized rates were suppressed due to both a very small underlying count plus extremely high variability. Confidence intervals can also be used to determine whether a rate in one health region is statistically below, above or no different than the rate for the same indicator in another health region.
- ◆ The confidence intervals for the age-standardized cancer incidence rates were produced via the Spiegelman method (refer to section 2.2.1).
- ◆ Cancer incidence rates are age-standardized using the direct method and the 1991 Canadian Census population structure. See “Age-standardized rates” section for details.
- ◆ Cancer incidence data were assigned to health regions using postal codes reported with place of residence and the automated geo-coding system (*PCCF+*) developed in Health Statistics Division. Where possible, remaining cancer incidence data (for which there were no postal codes available) were linked to health regions using the census subdivision (CSD) of residence.

## 4. Indicators Based on Statistics Canada Surveys

### 4.1 National Population Health Survey

The National Population Health Survey (NPHS), which began in 1994/95, collects information about the health of the Canadian population every two years. It covers household and institutional residents in all provinces and territories, except persons living on Indian reserves, Canadian Forces bases, and in some remote areas. The NPHS has both a longitudinal and a cross-sectional component. Respondents who are part of the longitudinal component will be followed for up to 20 years.

The *Health Indicators* data are based on both the longitudinal and cross-sectional components for household residents (institutional excluded) living in the provinces (territories excluded). Data are available for the first three cycles (1994/95, 1996/97 and 1998/99).

The 1994/95 and 1996/97 cross-sectional samples are made up of longitudinal respondents and their household members and individuals who were selected as part of supplemental samples, or "buy-ins", in some provinces. The 1998/99 cross-sectional sample is made up mostly of longitudinal respondents and their cohabitants. No buy-ins were added to 1998/99 data. However, to keep the sample representative, infants born in 1995 and thereafter and immigrants who entered Canada since the beginning of 1995 were randomly selected and added to the NPHS sample.

The 1994/95 provincial, non-institutional **cross-sectional** sample consisted of 27,263 households, of which 88.7% agreed to participate in the survey. After application of a screening rule to maintain the representativeness of the sample, 20,725 households remained in scope. In 18,342 of these households, the selected person was aged 12 or older. Their response rate to the in-depth health questions was 96.1% or 17,626 respondents. In 1996/97, the overall response rate at the household level was 82.6%. The response rate for the randomly selected individuals aged 2 or older in these households was 95.6%. A total of 81,804 respondents answered the in depth health questions in 1996/97. In 1998/99, the overall response rate was 88.2% at the household level. The response rate for the randomly selected respondents 0 or older in these households was 98.5%. A total of 17,244 respondents answered the in depth health questions in 1998/99.

The 1994/95 provincial, non-institutional **longitudinal** sample consisted of 17,276 respondents. A response rate of 93.6% was achieved in 1996/97, and a response rate of 88.9% was achieved in 1998/99.

### 4.2 National Population Health Survey -- Northern Component

Statistics Canada conducted the northern component of the NPHS in conjunction with the statistical bureaus in Yukon and NWT. Data were obtained through a separate survey due to the special challenges of survey taking in Canada's North.

The target population of the Yukon/NWT integrated NPHS/NLSCY survey included household residents living in private occupied dwellings located in the two territories, with the exclusion of populations on Indian Reserves, Canadian Forces Bases and in institutions. Moreover, persons

living in unorganized areas in the Yukon (13% of the population) and persons living in unorganized areas, very small or extreme northern communities of the NWT (4.9% of the population) were also excluded from the target population.

Most of the core content from the 1994-95 NPHS main survey is included in the northern survey; however, special "focus content" on stress was excluded. In each selected household in the North, demographic information was collected from all household members, then one person, aged 12 years and over, was randomly selected for a more in-depth interview. The questionnaire included components on health status, use of health services, risk factors and demographic and socio-economic status. Some content changes were made in the 1996/97 NPHS North survey.

Collection operations ran from November 1994 to March 1995 (and again from November 1996 to March 1997). Computer-assisted personal interviewing (CAPI), used for the NPHS in the provinces, was not available in the territories at the time of the survey. A paper and pencil questionnaire designed to replicate the CAPI application was used instead. Telephone interviews were conducted where available, otherwise personal interviews were done.

The selected person response rate for the NPHS 1994/95 was 94.2% at the North level (2,020 respondents). For the Yukon this rate was 94.8%, while the rate for the NWT was 93.1%. The cross-sectional response rate at the North level (both territories) for the NPHS 1996/97 was 86.2% (1,499 respondents). For the Yukon, this rate was 83.9% while the rate for the NWT was 89.8%.

**Heavy drinking, 1994/95:** Due to a high proportion (42.8%) of refusals/non-stated responses to the question on frequency of heavy drinking in the 1994/95 NPHS-North, these data were deemed unreleasable/unreliable. Heavy drinking has been defined as the number of times current drinkers drank 5 or more alcoholic beverages on one occasion.

For more information about the NPHS (Household or North component), see <http://www.statcan.ca/english/concepts/hs/index.htm>

### **4.3 Canadian Community Health Survey**

Health region level rates from the latest survey cycle (2003) are based on the boundaries in effect as of June 2003.

Starting with data year 2000/01, the Canadian Community Health Survey (CCHS) replaces the cross-sectional aspect of the NPHS.

The primary objective of the CCHS is to provide timely cross-sectional estimates of health determinants, health status and health system utilization at a sub-provincial level (health region or combination of health regions).

The CCHS collects information from individuals aged 12 or older who are living in private dwellings. People living on Indian reserves or Crown lands, residents of institutions, full-time members of the Canadian Armed Forces, and residents of certain remote regions are excluded. The CCHS covers approximately 98% of the Canadian population aged 12 or older.

Each two-year collection cycle is comprised of two distinct surveys: a health region-level survey in the first year with a total sample of 130,000 and a provincial-level survey in the second year with a total sample of 30,000. Sample sizes in any particular month or year may increase due to provincial or health region-level sample buy-ins.

The response rate for the first cycle of the CCHS at the national level was 84.7% (131,535 respondents). The response rate for the second cycle of the CCHS at the national level was 80.6% (135,573 respondents).

For more information about the CCHS, see <http://www.statcan.ca/english/concepts/hs/index.htm>.

### **4.4 National Longitudinal Survey of Children and Youth**

The National Longitudinal Survey of Children and Youth (NLSCY), developed jointly by Human Resources Development Canada and Statistics Canada, is a comprehensive survey which follows the development of children in Canada and paints a picture of their lives. The survey monitors children's development and measures the incidence of various factors that influence their development, both positively and negatively.

The first cycle of the NLSCY, which was conducted in late 1994 and early 1995, interviewed parents of approximately 23,000 children up to the age of 11. They shared information not only about their children, but also about themselves and the children's families, schools and neighbourhoods.

The second cycle, carried out in winter and spring of 1996-97, interviewed parents of the same children and provides unique insights into the evolution of children and their family environments over a two-year period. A new sample of newborn and 1-year-old children was added to cycle 2 to allow for cross-sectional estimates.

Collection of cycle 3 began in the fall of 1998 and was carried until June 1999. In addition to the original sample of children, who were aged 2 to 13 years at the time of the second data collection, a new sample of newborn and 1-year-old children was added to cycle 3 to allow for cross-sectional estimates. An extra cross-sectional sample of children 5 years old was also added to allow some provincial estimates for that age group.

#### **4.5 Bootstrapping**

To ensure high data quality for estimates from the NPHS, the CCHS and NLSCY, a weighted bootstrap resampling procedure (and for the NPHS-North, a modified bootstrap procedure) was used to calculate coefficients of variation (CVs) for totals and rates. If the CV was greater than 33.3% or the sample size was less than 10, the data were suppressed and an 'F' symbol appears in the data cell. If the CV is greater than 16.5% and no greater than 33.3%, the data should be interpreted with caution and an 'E' symbol appears in the same cell as the data. Data with CVs of 16.5% or less are presented without restrictions.

Sampling theory dictates that sample survey results of exactly 100% or 0% must have a coefficient of variation of exactly 0. In reality it is possible that in rare circumstances the true estimate may be lower than 100% or conversely greater than 0% and results should be interpreted as such.

### **5. Indicators Based on Crime Data (Statistics Canada)**

- ◆ Health region level data are not available for the crime-related indicators.
- ◆ Data on crime incidents that come to the attention of the police are captured and forwarded to the Canadian Centre for Justice Statistics (CCJS) via the Uniform Crime Reporting (UCR) Survey according to a nationally-approved set of common scoring rules, categories and definitions. The survey has been in operation since 1962 and has full national coverage.
- ◆ The UCR is a summary or aggregate-based survey that records the number of criminal incidents reported to the police. The survey does not gather information on the victims, but does collect information on the number of persons charged by sex and by an adult/youth breakdown. For all violent crimes (except robbery), a separate incident is counted for each victim. For non-violent crimes, one incident is counted for each distinct occurrence. Incidents that involve more than one infraction are counted under the most serious violation. As a result, less serious offences are under-represented by the UCR survey.
- ◆ The aggregate UCR survey scores violent incidents (except robbery) differently from other types of crime. For violent crime, a separate incident is recorded for each victim (i.e. if one person assaults three people, then three incidents are recorded; but if three people assault one person, only one incident is recorded). Robbery, however, is counted as if it were a non-violent crime in order to avoid inflating the number of victims (e.g. for a bank robbery, counting everyone present in the bank would result in an over-counting of robbery incidents).
- ◆ The aggregate UCR Survey records the total number of youths (aged 12 to 17) and adults (aged 18 and over) charged. When a person is charged with more than one offence, they are

counted only once, under the most serious offence. The most serious offence is generally the offence that carries the longest maximum sentence under the Criminal Code of Canada. Violent offences always take precedence over non-violent offences.

- ◆ The comparison between youth and adult crime rates poses some difficulties. The entire youth population represents a high-risk group for becoming involved in criminal activity. By contrast, the level of risk among adults is not consistent across the entire age group. Almost half of the adult population is 45 years and older; this age group is affected by fewer risk factors and as a result, is rarely involved in crime. A more direct comparison would look at youths and young adults. Unfortunately, data are not currently available to make this comparison.
- ◆ With UCR charge data it is possible for someone to be charged (and counted) more than once in a year. As a result, it is likely that the actual number of persons charged is less than the figure reported for a given time period.
- ◆ Rates are calculated on the basis of 100,000 population.

## **6. Indicators Based on Labour Force Data (Statistics Canada)**

Latest health region level rates are based on the boundaries in effect as of June 2005.

- ◆ Regional unemployment rates and youth unemployment rates were calculated as annual averages from the Canadian Labour Force Survey (LFS). The estimates were derived by linking, at the enumeration area (EA) level, the LFS geography to health regions.
- ◆ The LFS is a monthly sample of approximately 52,000 households. The survey is designed to represent the Canadian population aged 15 years and older. The survey excludes Indian reserves, full time members of the Canadian Forces, and persons living in institutions. The survey also excludes the Territories.
- ◆ The areas that are excluded from the LFS affect estimates for Peer Groups F and H. Just over 40% of the population of Peer Group F is excluded, while less than 10% of Peer Group H is excluded. As a result, estimates for Peer Group F are not available.
- ◆ Some health regions could not be published as the estimated rate did not meet the minimum requirements for quality and confidentiality.
- ◆ The unemployment rate is the number of unemployed persons divided by the labour force population, expressed as a percentage.

- ◆ An unemployed person is someone who:
  - was without work and had looked for work; or
  - was on temporary layoff and available for work; or
  - had a new job to start in the future.
- ◆ The labour force population consists of the unemployed people plus the employed persons. To be employed, a person
  - worked at any job at all; or
  - had a job but was not at work during the reference week

## **7. Indicators Based on Census Data (Statistics Canada)**

Health region level rates are based on the boundaries in effect as of June 2003 unless otherwise noted.

- ◆ Regional data on non-medical determinants of health and certain community characteristics were extracted from the Census, based on enumeration areas (EA) (for the 1996 Census) and census blocks (CBs) and Dissemination Areas (DA) (for the 2001 Census). A correspondence file, linking the Census geography to current health regions has been developed by the Health Statistics Division of Statistics Canada with the cooperation of provincial Ministries of Health, Alberta Treasury and BC Stats.
- ◆ Income-related indicators from any Census are based on income from the previous year.
- ◆ Low income rate, children in low income families: Low income data were not derived for economic families or unattached individuals in the Territories or on Indian reserves. For health regions containing Indian reserves, analysis of low income data should only be done with this caveat explicitly noted.
- ◆ Housing affordability: Farm homes and band housing on Indian reserves were not included in the calculation of housing affordability. For health regions containing Indian reserves, analysis of housing affordability should only be done with this caveat explicitly noted.
- ◆ Proportion Aboriginal population: This variable is derived from three questions asked in the Census (20% sample). Aboriginal population refers to those persons who reported identifying with at least one Aboriginal group, i.e. North American Indian, Métis or Inuit and/or those who reported being a treaty Indian or a Registered Indian as defined by the Indian Act of Canada and/or who were members of an Indian Band or First Nation. Census coverage studies were used to adjust these data with the population estimates for incompletely enumerated Indian Bands or reserves. The STC demographic population

estimates (which adjusts for census under coverage and refusal reserves) for the particular Census year were used as the denominator for these percentages.

- ◆ Owner-occupied dwellings: Band housing on Indian reserves and collective dwellings were not included. For health regions containing Indian reserves, analysis of owner-occupied dwellings should only be done with this caveat explicitly noted.
- ◆ Average expected dwelling values: The same exclusions for owner-occupied dwellings apply for this variable, in addition to farms.

**For more information on census concepts, please refer to the *2001 Census Dictionary, 92-378-XIE*.**

## **8. Health System Indicators (Canadian Institute for Health Information - CIHI)**

- CIHI's Privacy and Confidentiality policy does not permit the publication of data that might reasonably identify an individual, whether a patient or care provider, without consent. As a result, measures were taken to protect against residual disclosure from the dissemination of the regional rates including the suppression of small cell sizes. In addition, reporting data based on the region of the patient's residence (not hospitalization) reduces opportunities for identifying individual care providers.

### **8.1 Hospitalization Data and Rates (CIHI)**

- ◆ Data are reported based on the region of the patient's residence, not region of hospitalization. Consequently, these figures reflect the hospitalization experience of residents of the region wherever they are treated, as opposed to the comprehensive activity of the region's hospitals (who will also treat people from outside of the region).
- ◆ Regional rates for British Columbia were derived from the reported postal codes and a translation file developed by BC STATS, BC Ministry of Finance and Corporate Relations. Health region level data for the other provinces and territories were produced through a geo-coding process using correspondence files developed with input from each provincial and territorial health ministry and Alberta Treasury. The link between dissemination areas and health regions was first created to provide the best resolution to census geography, and a census subdivision link to health regions was derived from this file. The boundaries are those that were in effect as of April 2005. Records with invalid, missing, or partial postal codes are not included in the regional rates. The absence of complete postal codes from Quebec may affect rates for the Champlain LHIN (Ottawa area) and other border regions.

- ◆ Where possible, an all-Canada rate is provided for comparison purposes.
- ◆ At the Canada level and provincial levels, rates for health data that are based on a fiscal year (April to March) use October 1<sup>st</sup> population estimates. Unless otherwise specified, Canadian and provincial hospitalization rates are standardized using the same methodology as regional rates. Other rates are based on appropriate population figures. Standardized rates are age-adjusted using a direct method of standardization based on the July 1<sup>st</sup>, 1991 Canadian population. See section 2.2.2 for details.
- ◆ Unless otherwise specified, hospitalizations include discharges and deaths for inpatients in acute care hospitals for the reference period. Same day surgery (outpatient) cases and patients admitted to non-acute care hospitals (e.g. chronic care, psychiatric or rehabilitation facilities) are not included in the totals.
- ◆ Indicators based on the Discharge Abstract Database include only jurisdictions that submit comprehensively to the database. Therefore, data from Quebec and some parts of Manitoba are not available.
- ◆ ICD-10-CA and the Canadian Classification of Health Interventions (CCI) systems of coding diagnoses and procedures came into effect April 1, 2001, and were adopted by Newfoundland, PEI, Nova Scotia, parts of Saskatchewan, British Columbia and the Yukon Territory. Ontario, the remainder of Saskatchewan, Alberta, the NWT and Nunavut implemented ICD-10-CA/CCI on April 1, 2002 and New Brunswick on April 1, 2003. Indicator cases that were originally coded in ICD-10-CA or CCI have been extracted on the relevant codes and not the ICD-9 or CCP translations.
- ◆ New and revised coding standards introduced with the ICD-10-CA/CCI classification systems may affect the comparability of rates with those appearing in previous releases for some of the indicators.
- ◆ Bypass Surgery: Variations in the use of this procedure may be related to utilization rates of coronary angioplasty, an alternative intervention for improving blood flow to the heart.
- ◆ The definition of the ambulatory care sensitive conditions (ACSC) indicator has been revised. Rates are not comparable to those published by CIHI in previous years. See Definitions for further information.
- ◆ Beginning with data year 2002/03, stillbirths are included in the Caesarean section rate. This change may affect the comparability of rates with those appearing in previous releases. See Definitions for further information.
- ◆ The methodology for calculating the acute myocardial infarction (AMI) readmission and asthma readmission indicators, as well as the codes used to extract the in-hospital hip fracture indicator have been revised. These changes may affect the comparability of rates with those appearing in previous releases. See Technical Notes for further information.
- ◆ The 30-day in-hospital mortality, in-hospital hip fracture, and readmission indicator rates are based on a three-year average. Due to differences in the way data are collected, these

indicators are not available for all provinces and territories. Therefore, the average (Canada) rate does not include all provinces/territories. [Technical notes and model specifications](#) are available for AMI and Stroke 30-day mortality, in-hospital hip fracture, as well as the readmission indicators (AMI, asthma, hysterectomy, and prostatectomy).

- ◆ Where information is available, cancelled, previous, and “abandoned after onset” procedures are excluded from the calculations. For Quebec data, cancelled procedures are not reported and therefore have not been excluded.

## 8.2 Physician Data (CIHI)

While physician density ratios are useful indicators of changes in physician numbers relative to the population, inference from total numbers or ratios as to the adequacy of provider resources should not be made. Various factors influence whether the supply of physicians is appropriate, such as: distribution and location of physicians within a region or province; physician type (i.e., family medicine physicians vs. specialists); level of service provided (full-time vs. part-time); physician age and gender; population's access to hospitals, health care facilities, technology and other types of health care providers; population needs (demographic characteristics and health problems); and society's perceptions and expectations.

Physician counts include all active general practitioners, family practitioners and specialist physicians as of December 31 of the reference year. The data include physicians in clinical and non-clinical practice and excludes residents and physicians who are not licensed to provide clinical practice and have requested to the Business Information Group that their data not be published. For purposes of reporting, physician specialty classification is based on postgraduate certification credentials achieved in Canada. Physicians designated as family practitioners include certificants of the College of Family Physicians of Canada (CCFP and CCFP-Emergency Medicine). Specialist physicians include certificants of the Royal College of Physicians and Surgeons of Canada and/or the College des médecins du Québec. All other physicians, including non-College of Family Physicians general practitioners, foreign-certified specialists and other non-certified specialists, are included in the family practice counts. It is recognized that physician classification in this manner does not necessarily reflect the services provided by individual physicians. The range of services provided by a physician is subject to provincial licensure rules, medical service plan payment arrangements and individual practice choices. Therefore, rates may differ from other publications. For example, the Newfoundland Medical Board (NMB) grants full and provisional licenses to non-certified specialists. These physicians are counted as family practitioners in the *Health Indicators* rates, but are counted as specialist physicians in annual reports of the NMB Registrar.

## 8.3 National Health Expenditure Database (CIHI)

- ◆ Expenditure figures include spending by both the public and private sectors. For further information, see National Health Expenditure Trends, 1975-2004.

- ◆ Provincial per capita figures are affected by numerous factors that will affect inter-provincial comparisons including, but not limited to, differing provincial inflation rates that are related to provincial differences in arbitration agreements between provincial governments and, for example, medical associations; different population distributions; geography; and differences in provincial purchasing power.

## **Appendix**

### **Population estimates**

In summary, the population estimates for the health regions in this publication were prepared as follows. First, the 2001 population estimates were based on the 2001 Census, adjusted for net undercoverage. Secondly, for non-census years, the Census-based population estimates were adjusted for changes in the population, primarily using administrative data. Lastly, subprovincial/subterritorial population estimates were controlled to sum to the population estimates at the provincial/territorial level prepared by Statistics Canada.

### **Statistics Canada Methodology**

For health regions in all provinces/territories except for Alberta and British Columbia, the method is that of Statistics Canada, described below. These estimates are based on census data and on population estimates by census division (CD) produced by Demography Division.

#### **Population universe**

The population included in these estimates is based on the 1996 and 2001 Census of Canada. The universe, therefore, corresponds to the census universe, which includes Canadian citizens and landed immigrants with a usual place of residence in Canada (or abroad on a Canadian Forces base, attached to a diplomatic mission or aboard merchant vessels) and non-permanent residents. For a complete description of the census universe, refer to the *2001 Census Dictionary*, Statistics Canada Catalogue no 92-378-XPE.

Census data are adjusted to take into account net census undercoverage and incompletely enumerated Indian reserves. Components of population change occurring between census day and July 1, 1996 or 2001 were also taken into account.

#### **Method for deriving demographic estimates by health regions**

Population estimates at the census division (CD) level are produced annually by Demography Division, using the component method, which accounts for changes in the number of births and deaths, as well as intra-provincial, inter-provincial and international migration. The CD population estimates are prorated to the provincial population estimates.

Demographic estimates by health region are derived from these estimates using the following steps:

***Calculation of conversion factors:***

Health Statistics Division has created a file linking dissemination areas (DAs), census subdivisions (CSDs) and census divisions (CDs) for each health region.

The 2001 Census population by CDs (DAs or CSDs in case of split CDs) was adjusted for late changes in geography and for partially enumerated Indian reserves to ensure DAs added up to the CD.

In cases where health regions split CDs, the proportion of population in each CD split is used to derive 'conversion factors' in order to allocate the CD level estimate to the appropriate health region.

***Application of conversion factors to CD demographic estimates:***

Conversion factors were applied to CD population estimates for 1996 to 2001 allowing conversion of CD population estimates into HR population estimates. The same approach used to obtain the total population estimates was used to establish age and sex estimates by HR.

**Evaluation of the method**

To assess the quality of demographic estimates by HR, two evaluations were performed. First, an estimate was produced for July 1, 2001 with a postcensal CD demographic estimate based on the *2001 Census* which was compared with the 2001 Census counts. This comparison indicated that the conversion factors were providing results comparable with the census counts by HR.

Second, the same method was also used to generate estimates for 2001 by HR from a postcensal CD estimate based on the *1996 Census*. The results were compared with the actual 2001 counts as derived from the 2001 Census, thereby allowing an evaluation of the accuracy of the method over a five-year period. The results indicated that the methodology adequately accounted for the demographic changes.

For more information on the method used to produce health region or CD estimates, contact Hubert Denis of the Demography Division of Statistics Canada at (613) 951-2327 or e-mail Hubert Denis @statcan.ca.

## Provincial Methodologies

For health regions in Quebec, Alberta and British Columbia, the administrative files used for adjusting the 2001-based population estimates for non-census years differ from those used by Statistics Canada in the description above. For these provinces, the health region population estimates were supplied by Alberta Health and Wellness, and BC Stats.

### Alberta

Statistics Canada used population estimates by health regions prepared by the Alberta Health and Wellness department. These estimates are derived from the Alberta Stakeholder Register File containing information on individuals who are entitled to basic medical services under the Alberta Health Care Insurance Plan (AHCIP). Population growth as derived from this file for each of Alberta's health regions was then applied to Demography Division's estimates for 2001 to account for annual changes in the population. The estimates were further adjusted to correspond to the provincial estimates produced annually by Demography Division.

For more information on Alberta's estimates, refer to *Calculating Demographic and Epidemiological Quantities in Alberta by Geo-Political Areas, Health Surveillance, Alberta Health and Wellness, March 2001* or contact Donald Schopflocher of the Health Strategies Division, Health Surveillance in Alberta at (780)422-4630 or e-mail at [donald.schopflocher@health.gov.ab.ca](mailto:donald.schopflocher@health.gov.ab.ca).

## British Columbia

The methodology used by British Columbia to derive small area populations by gender and age group is divided into two parts.

A Regression Approach, specifically the **Difference-Correlation Method (DCM)**, is the primary method underlying the sub-provincial population estimates. A secondary method, known as **Proportional Allocation (PA)**, is also used to estimate the population for certain classes of areas. Both these methods use information derived from a set of indicators obtained from administrative files that are symptomatic of regional population changes.

In essence, the British Columbia small area population estimation model works as follows. Beginning with the most recent federal census (in this case the 2001 Census of Canada), each region's share of provincial population is adjusted up or down according to the current share of the provincial total of a weighted combination of residential hydro connections and/or Old Age Security recipients. Estimates of the population living in municipalities along with that portion of the population living outside the municipality but within the regional district (i.e., unorganized area), are controlled at the provincial level by a British Columbia population estimate prepared by Statistics Canada. Regional district population estimates are derived by summing the municipal and unorganized area population estimates. Local health areas are also controlled at the provincial level, and in order to ensure consistency, the local health area population estimates within each regional district are then tied to the regional district population estimates.<sup>13</sup>

---

13. *BC Stats. Generalized Estimation System (GES), Small Area Population Estimation Methodology.* Ministry of Finance and Corporate Relations, December 1998.

Detailed information about the methodology used for estimating the age/gender distribution of small area population in British Columbia can be found in two documents on the BC Stats website [www.bcstats.gov.bc.ca](http://www.bcstats.gov.bc.ca) "Generalized Estimation System (GES)", December 1998, and "Estimating the Age/Gender Distribution of Small Area Populations in British Columbia", April 1994.