The Impact of Estimation Method and Population Adjustment on Canadian Life Table Estimates

Edward Ng and Jane F. Gentleman*

Abstract

Abridged life tables centred on 1991 were produced from the 1991 Canadian census, net census undercoverage estimates, and death data from 1990 to 1992. The sensitivity of life table values to differing methods of estimation and population estimates was investigated. The results from four methods by Greville, Chiang, and Keyfitz were compared, and population estimates, both adjusted and unadjusted for net census undercoverage, were used to test the effects of method and type of population estimate on life table values. The results indicate that the method used to derive the estimates had much less influence on the life table values than did the choice of population estimate.

The change to life expectancy at birth due to the method of calculation chosen was at most 15 days, whereas the change due to the population estimate chosen was about 73 days. Since there are age, sex and provincial variations in net undercoverage rates, life expectancies differed accordingly.

Keywords: *abridged life tables, Greville, Keyfitz, Chiang, census undercoverage*

Introduction

Life expectancy is one of the most widely used population health indicators. These estimates are used by public health workers, demographers, actuaries, policy makers, and many others in studies of longevity, fertility, migration, and population growth. Estimates of life expectancy are derived from a demographic model that describes and summarizes the mortality and survivorship of a population in the form of a life table.

There are two kinds of life table: *cohort* and *current*.[2] Cohort life tables (which require longitudinal data) are based on the mortality rates of a group of individuals tracked over time. Consequently, a considerable period is required for data collection. Current (or period) life tables (which require cross-sectional data) describe a hypothetical cohort of 100,000 individuals subjected to current age-sex-specific mortality rates. Current life tables, which are of interest here, can be considered as describing a stationary population—that is, a population with a fixed age-sex structure and a zero rate of increase, which would be the long-term result of unchanging age-sex-specific mortality rates and a constant annual number of births.

Note to readers

Results in this article are based on abridged life tables, using five-year age groups, and differ slightly from results in **Life Expectancy of Canadians** (in this issue) that are based on detailed life tables, using one-year age groups.

Life expectancy and other life table elements may be calculated for single years of age if the population is sufficiently large. The Health Statistics Division of Statistics Canada publishes *detailed* life tables (by single year of age) for nine provinces (excluding Prince Edward Island and the territories) and for Canada (including Prince Edward Island and the territories). These tables are produced every five years using death counts for three years centred on a census year to smooth the estimates.[3-5]

---

* *Edward Ng (613-951-1733) and Jane Gentleman (613-951-8553) are with the Health Statistics Division at Statistics Canada, Ottawa K1A 0T6. An earlier version of this paper was presented at the 1994 Joint Statistical Meetings in Toronto.[1]*

At the same time, *abridged* life tables (by five-year age group) are produced for all provinces and Canada using death and population data from the same years.

Population counts are among the data needed to produce life tables. Since 1993, Statistics Canada has produced annual population estimates that account for net census undercoverage, which is the difference between census undercoverage and census overcoverage. Undercoverage refers to persons who were not enumerated in the census but should have been, whereas overcoverage refers to persons counted more than once, or persons who were counted but should not have been (such as diplomats or visitors).[6]

The series of adjusted population counts was produced retroactively to 1971. Once the 1991 census had been conducted, net census undercoverage evaluated,[7] and 1990-1992 death data compiled, the official Statistics Canada life tables for 1990-1992, which used adjusted population counts for the first time, could be produced.[4]

Practically, life expectancy depends on health status, age, sex, marital status, lifestyle, income, place of residence, etc. Statistically, life expectancy is influenced by the method used to calculate life tables and by the population estimates chosen. This study investigates the effects on current Canadian abridged life tables of using four different methods of calculation and adjusting the population counts for net undercoverage. Comparisons are made by age, sex, and province, because the amount of undercoverage adjustment varies considerably for each factor.[6,7]

Data and methods

Abridged life table construction methods attributable to Greville,[8] Chiang,[9,10] and Keyfitz (an iterative method[11-14] and a non-iterative method[15,16]) were examined. One basic difference among the four methods is how they calculate mortality rates from death and population counts (see *Appendix*). To construct the various current life tables for Canada[a] and the provinces, the following data are required:

- number of deaths in specific years, by age group and sex, for Canada and the provinces;
- number of births in specific years, by sex, for Canada and the provinces;
- population counts from the census, by age group and sex, for Canada and the provinces;
- the same population counts as above, adjusted for net census undercoverage.

A typical life table displays some of these data plus derived information, including death rates, mortality rates, and life expectancies (Table 1)(see *Life table elements*).

Life table elements

l_x = number alive at age x (out of an initial cohort of l_0 = 100,000)

$_nd_x$ = number in the cohort who died during the age interval from x to x+n

$_np_x$ = proportion of the cohort alive at age x who survived to the end of the age interval (x+n)

$_nq_x$ = 1 - $_np_x$, proportion of the cohort alive at age x who died during the age interval from x to x+n (mortality rate)

$_nL_x$ = number of person-years lived by the cohort during the age interval from x to x+n

T_x = number of person-years lived by the cohort at age x and all subsequent ages

e_x = life expectancy at age x (average number of years of life remaining)

$_nP_x$ = observed population count for the age interval from x to x+n

$_nD_x$ = observed number of deaths over 3 years centred on the census year, for the age interval x to x+n

$_nm_x$ = observed death rate for the age interval from x to x+n

Counts of deaths and births were obtained from the Canadian vital statistics database maintained by the Health Statistics Division of Statistics Canada. Adjustments for net undercoverage were made to census data by the Demography Division of Statistics Canada. In general, undercoverage exceeds overcoverage, so the adjusted population counts are larger than the unadjusted counts. Consequently, when the derivations are based on adjusted counts, mortality rates are lower and life expectancies are higher.

---

[a] *The life table for all of Canada includes the death and population data for the two territories.*

Table 1

Abridged life table, Canada, 1990-1992

Age group (years)	l_x	${}_nd_x$	${}_np_x$	${}_nq_x$	${}_nL_x$	T_x	e_x	${}_nP_x$	${}_nD_x$	${}_nm_x$
< 1	100,000	645	0.993552	0.006448	99,437	7,779,851	77.80	401,731	7,770	0.006447
1-4	99,355	132	0.998674	0.001326	397,174	7,680,414	77.30	1,551,438	1,544	0.000332
5-9	99,223	95	0.999039	0.000961	495,897	7,283,240	73.40	1,952,910	1,127	0.000192
10-14	99,128	105	0.998944	0.001056	495,423	6,787,343	68.47	1,912,988	1,213	0.000211
15-19	99,023	311	0.996857	0.003143	494,394	6,291,919	63.54	1,925,926	3,637	0.000629
20-24	98,712	371	0.996238	0.003762	492,648	5,797,525	58.73	2,108,995	4,769	0.000754
25-29	98,341	388	0.996058	0.003942	490,752	5,304,877	53.94	2,528,685	5,992	0.000790
30-34	97,953	454	0.995367	0.004633	488,674	4,814,126	49.15	2,597,980	7,238	0.000929
35-39	97,499	597	0.993877	0.006123	486,076	4,325,451	44.36	2,344,684	8,638	0.001228
40-44	96,902	802	0.991720	0.008280	482,652	3,839,375	39.62	2,138,771	10,666	0.001662
45-49	96,100	1,300	0.986474	0.013526	477,516	3,356,724	34.93	1,674,125	13,672	0.002722
50-54	94,800	2,078	0.978075	0.021925	469,229	2,879,208	30.37	1,339,856	17,806	0.004430
55-59	92,722	3,337	0.964007	0.035993	455,907	2,409,979	25.99	1,238,381	27,198	0.007321
60-64	89,384	5,165	0.942221	0.057779	434,857	1,954,072	21.86	1,190,172	42,401	0.011875
65-69	84,220	7,404	0.912091	0.087909	403,664	1,519,215	18.04	1,084,556	59,661	0.018337
70-74	76,816	10,324	0.865597	0.134403	359,605	1,115,551	14.52	834,014	71,818	0.028704
75-79	66,492	13,815	0.792231	0.207769	299,214	755,946	11.37	622,230	86,158	0.046155
80-84	52,677	16,527	0.686262	0.313738	222,615	456,732	8.67	382,310	85,089	0.074188
85-89	36,150	16,444	0.545132	0.454868	138,821	234,117	6.48	192,414	68,375	0.118451
90+	19,707	19,707	0.000000	1.000000	95,296	95,296	4.84	95,466	59,225	0.206793

Note: *Estimates were calculated using Greville's method and adjusted population counts.*

Results

Changing the method of calculation has a small impact

Life tables were examined for all four methods and both types of population counts, by sex and province (Table 2). The impact of the choice of method on life expectancy was highly consistent across sex, province, and the population counts used. Life expectancy estimates derived from Chiang's method almost exactly reproduced the results from Greville's. For example, for males in Canada, life expectancy at birth differed by only about 0.01 of a year, or about 4 days.

Both of Keyfitz's methods, in attempting to correct for deviation from stationary population assumptions (see *Appendix*), produced life expectancies that were generally lower than those derived from Greville's and Chiang's methods. For both sexes, at birth and at age 65, Keyfitz's methods, in most cases, produced life expectancies of at most 0.04 of a year (15 days) lower than those calculated using the other two methods.

The lower life expectancies derived from Keyfitz's methods were attributable to differences in the method of calculating mortality rates, ${}_nq_x$. Beyond age 40, Keyfitz's mortality rates were consistently higher than those of Greville and Chiang, thus yielding lower life expectancies. However, differences among life expectancy estimates derived by different methods of calculation were small, compared with differences produced by the use of different population counts.

Adjusting for net census undercoverage has a larger impact

Life expectancy estimates based on adjusted population counts were, as expected, consistently higher than those based on unadjusted counts, regardless of the method used (Table 2). For example, life expectancy at birth was 0.20 of a year (73 days) longer for both sexes when adjusted population estimates were used (with Greville's method) instead of unadjusted counts; at age 65, the difference was 0.12 of a year (about 44 days).

Table 2

Life expectancy at birth and at age 65, by method of life table construction, type of population count, sex, and province, Canada, 1990-1992

	Greville		Chiang		Keyfitz's iterative		Keyfitz's non-iterative	
	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted
				Life expectancy at birth Years				
Canada - both sexes	77.80	77.60	77.80	77.60	77.76	77.56	77.76	77.57
Male	74.61	74.38	74.62	74.37	74.57	74.33	74.57	74.34
Female	80.97	80.81	80.97	80.80	80.93	80.77	80.94	80.78
Newfoundland	76.51	76.38	76.55	76.42	76.47	76.34	76.47	76.34
Prince Edward Island	76.86	76.80	76.87	76.80	76.84	76.78	76.85	76.79
Nova Scotia	77.00	76.88	77.02	76.89	76.97	76.85	76.98	76.86
New Brunswick	77.55	77.33	77.56	77.33	77.51	77.30	77.52	77.31
Quebec	77.40	77.20	77.41	77.20	77.35	77.15	77.36	77.16
Ontario	78.02	77.79	78.03	77.79	77.98	77.76	77.99	77.76
Manitoba	77.68	77.56	77.69	77.56	77.65	77.53	77.66	77.54
Saskatchewan	78.33	78.20	78.33	78.20	78.30	78.17	78.31	78.19
Alberta	78.07	77.93	78.07	77.92	78.02	77.89	78.03	77.89
British Columbia	78.28	78.08	78.27	78.07	78.24	78.04	78.25	78.05
				Life expectancy at age 65 Years				
Canada - both sexes	18.04	17.92	18.04	17.93	18.01	17.89	18.01	17.90
Male	15.80	15.70	15.81	15.70	15.77	15.66	15.77	15.66
Female	19.98	19.86	19.98	19.86	19.95	19.83	19.95	19.83
Newfoundland	16.69	16.61	16.70	16.62	16.66	16.58	16.66	16.58
Prince Edward Island	17.41	17.38	17.42	17.40	17.40	17.38	17.41	17.38
Nova Scotia	17.39	17.32	17.40	17.33	17.37	17.30	17.38	17.31
New Brunswick	17.91	17.79	17.92	17.80	17.89	17.77	17.90	17.78
Quebec	17.78	17.66	17.79	17.67	17.74	17.62	17.75	17.63
Ontario	18.01	17.87	18.01	17.88	17.98	17.84	17.98	17.84
Manitoba	18.08	18.01	18.08	18.02	18.06	17.99	18.06	18.00
Saskatchewan	18.70	18.63	18.70	18.63	18.68	18.62	18.69	18.62
Alberta	18.47	18.39	18.47	18.40	18.44	18.36	18.44	18.36
British Columbia	18.59	18.48	18.59	18.48	18.56	18.45	18.56	18.45

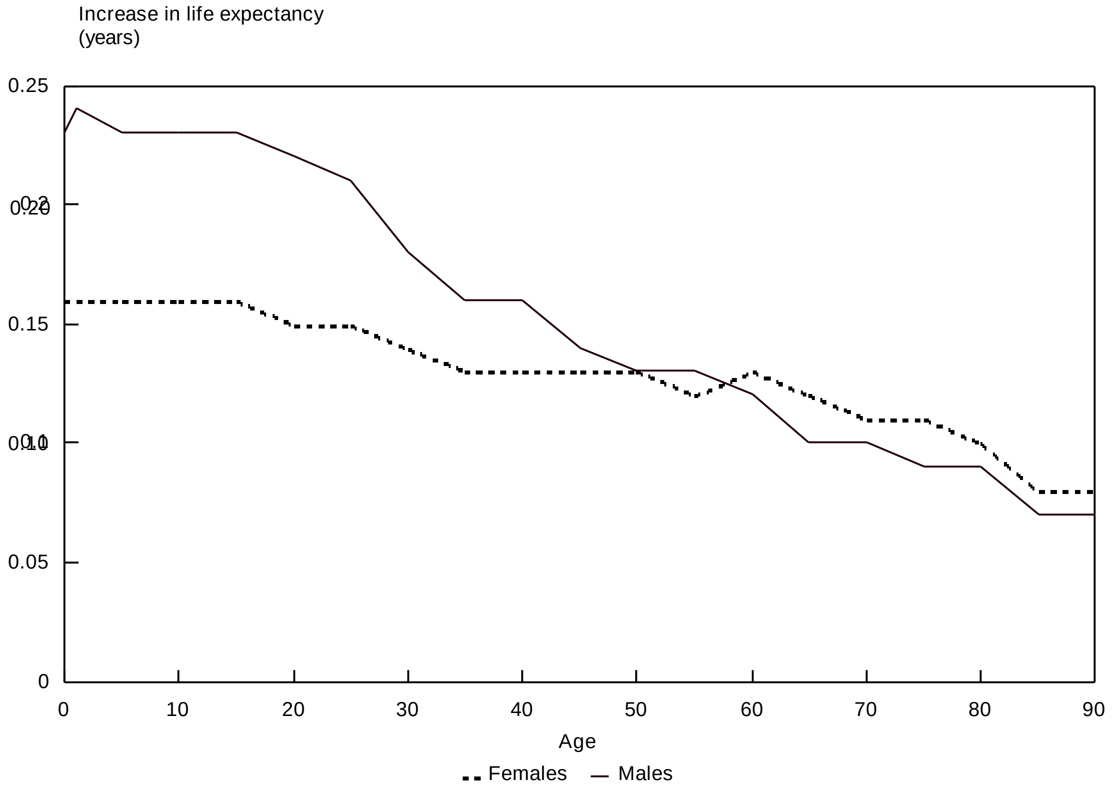
Note: *Life expectancy estimates for Newfoundland and Prince Edward Island should be considered with caution because of the small number of deaths and the small population in these provinces.*

Since there were age, sex and provincial differences in net undercoverage rates,[7] life expectancies differed accordingly. In general, net undercoverage was higher for males than for females (3.4% and 2.4%, respectively). For males, the highest net undercoverage rates were at ages 20 to 24 and 25 to 34: 7.8% and 6.8%, respectively. The pattern for females was similar, with corresponding rates of 6.3% and 3.5%. The increase in life expectancy resulting from using adjusted counts was higher for males at younger ages and higher for females at older ages. This crossover reflected the higher net undercoverage among men aged 20 to 64 and the lower rates among men aged 65 and over, compared with their female counterparts (Chart 1). At age 65 and over, the net undercoverage rates were 0.8% for men and 1.0% for women.

Consequently, the use of adjusted instead of unadjusted counts (with Greville's method) increased life expectancy at birth by 0.23 of a year (84 days) for males and 0.16 of a year (58 days) for females. At age 65, the difference was 0.10 of a year (37 days) for men and 0.12 of a year (44 days) for women.

Chart 1

Increase in life expectancy estimates due to adjusted population counts, by sex, Canada, 1990-92



Note: *Life expectancy estimates were calculated using Greville's method.*

Without population adjustment, the excess life expectancy of females over males at birth was 6.43 years; with population adjustment, the excess was reduced by 0.07 of a year to 6.36 years. However, at age 65, the excess life expectancy of women over men was 4.16 years without population adjustment and 4.18 years with adjustment, an *increase* of 0.02 of a year.

The highest provincial net census undercoverage rate was in Ontario (3.6%), and the lowest, in Prince Edward Island (0.9%).[7] Thus, the increase in life expectancy at birth for both sexes combined caused by adjusting the census counts was greatest for Ontario (about 0.23 of a year or 84 days) and least for Prince Edward Island (about 0.06 of a year or 22 days), based on Greville's method. At age 65, the increase in remaining life expectancy was 0.14 of a year (51 days) for Ontario and 0.03 of a year (11 days) for Prince Edward Island.

Discussion

The World Health Organization asserts that although different methods of calculating life tables (including the methods of Greville, Chiang, and Keyfitz) are based on very different assumptions, they do not yield differences significant to mortality research.[17,18] This finding was confirmed using each method and Canadian vital statistics from 1990-92 to derive abridged life tables.

In fact, the use of 1991 population counts adjusted for net census undercoverage had a much larger impact than the method of calculation. For Canada as a whole, the change to life expectancy at birth due to the method of calculation was 15 days at most, whereas the change due to the population estimates chosen was about 73 days.

Since the unadjusted population count was an underestimate for all age groups, use of the adjusted population consistently revised the mortality rates downward, yielding higher life expectancies. The impact of the population adjustment was sizeable even at older ages, at which net undercoverage was under 1% (compared with an overall undercoverage rate of 3%).[7] For example, at age 80, the difference was about 0.10 of a year (37 days) in remaining life expectancy. This sizeable upward adjustment in life expectancy at older ages can be explained by the higher number of life table survivors resulting from the lower mortality rates.

In Canada, where the registration of births and deaths is almost complete, accounting for census undercoverage in the construction of life tables is important because it is desirable to have the same degree of completeness of coverage for all data. Statistics Canada produces both unadjusted and adjusted population figures for census years. The Health Statistics Division of Statistics Canada uses the adjusted populations to calculate rates in its standard data products.

Using an adjusted population count to calculate a mortality rate gives a more exact measure of the population at risk of death. However, caution must be exercised for international comparisons, since in their life table construction, most other countries do not adjust their official population counts for net census undercoverage. Similar caution is required for historical comparisons in Canada, because all official life tables previously published by the Health Statistics Division (except for the most recent series for 1990-1992[4]) did not use population counts adjusted for net census undercoverage.

Appendix

Summary of four methods of abridged life table construction

One difference among the four methods of Greville, Chiang, and Keyfitz is the transformation from the observed death rate (${}_nm_x$) to the mortality rate (${}_nq_x$). The observed annual death rate is the ratio of the deaths, averaged over 1990-92 (${}_nD_x / 3$), to the population count (${}_nP_x$). The mortality rate relates deaths to the appropriate population exposed to the risk of death at the beginning of the age interval. Conventionally, equation 1 is used to convert from the observed death rate to the mortality rate for all ages other than infancy:

$$ {}_nq_x = \frac{{}_nm_x}{\frac{1}{n} + 0.5\ {}_nm_x} \qquad (1) $$

Equation 1 assumes that the deaths over n years are linearly distributed, so that by the mid-point of the period, half of the deaths have occurred. This assumption is an oversimplification, given that mortality generally increases non-linearly with age for adults.

For infants, birth data are also used to estimate the mortality rate. Since some deaths in a year occur to infants born in that year and some occur to those born in the previous year, a special procedure is used by Greville and Keyfitz to subdivide the deaths according to birth cohort. This special procedure is needed to better estimate the true infant population exposed to the risk of death.[2]

In 1943, Greville[8] improved the formula (equation 1) by assuming that the observed death rate (${}_nm_x$) is exponentially related to age (x) in accordance with the Gompertz law of mortality. In other words, there is a linear relationship between age (x) and the natural log of the observed death rate (ln ${}_nm_x$). Denoting the slope of this line as ln C, Greville's conversion of the observed death rates to mortality rates (${}_nq_x$) is:

$$ {}_nq_x = \frac{{}_nm_x}{\frac{1}{n} + {}_nm_x \left[0.5 + \frac{n}{12} \left({}_nm_x - \ln C \right) \right]} \qquad (2) $$

Note that equation 2 is equivalent to equation 1 if ${}_nm_x$ is the same as ln C. If it is not, there will be upward or downward adjustment of the resultant ${}_nq_x$, depending on whether ${}_nm_x$ is smaller or larger than ln C.

Instead of relying on the Gompertz law of mortality, Chiang[9,10] empirically derived the mean proportion of the last age interval of life lived (${}_na_x$) from death certificates for which the date of birth of the deceased was available. For example, when a person dies on his or her 33rd birthday, that person has lived 3/5 of the age interval between 30 and 34.

The value of ${}_na_x$ for this person is 0.6. The formula used by Chiang is:

$$ {}_nq_x = \frac{{}_nm_x}{\frac{1}{n} + (1 - {}_na_x)\,{}_nm_x} \qquad (3) $$

Equation 3 is equivalent to equation 1 if ${}_na_x$ is equal to 0.5.

A current life table can be considered as describing a stationary population. However, the population sizes of successive cohorts in most countries are not constant. The violation of the basic stationary assumption was the main motivation for Keyfitz's alternative approaches to life table construction.[11-16] Keyfitz's 1966 approach[11] is complex and can only be summarized here. He introduced a parameter, ${}_nr_x$, representing the age-specific rate of the population increase. The values of ${}_nr_x$ influence the distribution of population and death within the age groups. The more rapid the population growth, the lower the observed age-specific death rates will be. To infer the age-specific rates in the stationary population from the observed rates, corrections have to be made through successive iterations to adjust ${}_nq_x$ and thus the life expectancy. The iteration stops when the life table yields age-specific death rates as close to the death rates observed in the population as possible. Here, Keyfitz's equation to convert from the death rate (${}_nm_x$) to the mortality rate (${}_nq_x$) is the same as Chiang's (equation 3), except that the quantity ${}_na_x$ is derived and modified during the iterations.

Since the iterative method is relatively complicated to implement, in 1975 Keyfitz[15,16] developed a much simpler alterative that requires no iteration but gives equally precise results:

$$ {}_nq_x = 1 - e^{[-5\,({}_nm_x + K)]} \qquad (4) $$

where

$$ K = \frac{({}_nP_{x-5} - {}_nP_{x+5})\,({}_nm_{x+5} - {}_nm_{x-5})}{48\,{}_nP_x} \qquad (5) $$

This method also explicitly deals with the change in population (${}_nP_x$). Whenever the population is declining with age and the age-specific death rates are rising, K will be a positive value, and this will result in a higher mortality rate.

Acknowledgement

We thank Russell Wilkins and Jiajian Chen of the Health Statistics Division, Statistics Canada, for providing support and suggestions. Appreciation is also extended to Pierre David of the Household Survey Methods Division, Statistics Canada, for deriving the required values of Chiang's ${}_na_x$'s from the Canadian vital statistics data.

References

1. Ng E, Gentleman J. Life expectancy in Canada: an empirical comparison of different estimation methods. In: *Proceedings of the Social Statistics Section, American Statistical Association, Joint Statistical Meetings*: 1994 Aug 13-18; Toronto: 88-93.

2. Shryock HS, Siegel JS. *The methods and materials of demography*. Condensed edition, Stockwell EG. editor. San Diego: Academic Press, 1976.

3. Statistics Canada. Life tables, Canada and provinces, 1985-1987. *Health Reports* (Cat. No. 82-003S Suppl 13) 1991.

4. Statistics Canada. *Life tables, Canada and provinces, 1990-1992* (Cat. No. 84-537) Ottawa: Statistics Canada, 1995.

5. Millar W. Life expectancy of Canadians. *Health Reports* (Statistics Canada, Cat. No. 82-003) 1995; 7(3): 23-26.

6. Bender R. Impact of new population estimates on health and vital statistics. *Health Reports* (Statistics Canada, Cat. No. 82-003) 1995; 7(1): 7-18.

7. Statistics Canada. *Coverage: 1991 Census technical reports* (Cat. No. 92-341E) Ottawa: Statistics Canada, 1994.

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

9. Chiang CL. *Introduction to stochastic processes in biostatistics*. New York: Wiley, 1968.

10. Chiang CL. *The life table and its application*. Malabar, Fl.: Robert E. Krieger Publishing Co., 1984.

11. Keyfitz N. A life table that agrees with the data. Part I. *Journal of the American Statistical Association* 1966; 61(314): 305-12.

12. Keyfitz N. A life table that agrees with the data. Part II. *Journal of the American Statistical Association* 1968; 63(324): 1253-68.

13. Keyfitz N. *Introduction to the mathematics of population*. Reading, Mass.: Addision-Wesley, 1968.

14. Keyfitz N. Finding probabilities from observed rates or how to make a life table. *The American Statistician* 1970; 24: 28-33.

15. Keyfitz N, Frauenthal J. An improved life table method. *Biometrics* 1975; 31: 889-99.

16. Keyfitz N. *Applied mathematical demography*. New York: Springer, 1985.

17. World Health Organization. *Manual of mortality analysis*. Geneva: WHO, 1977.

18. Peron Y. La construction de tables de mortalité abrégees: comparaison de trois méthodes usuelles. *Population* 1971; 26(6): 1125-30.