Research Article

Variance models of the last age interval and their impact on life expectancy at subnational scales

Ernest Lo

Dan Vatnik

Andrea Benedetti

Robert Bourbeau

Contents

Variance models of the last age interval and their impact on life expectancy at subnational scales

Ernest Lo[1]

Dan Vatnik[2]

Andrea Benedetti[3]

Robert Bourbeau[4]

Abstract

BACKGROUND
The Chiang method is the most widely accepted standard for estimating life expectancy (LE) at subnational scales; it is the only method that provides an equation for the LE variance. However, the Chiang variance formula incorrectly omits the contribution of the last age interval. This error is largely unknown to practitioners, and its impact has not been rigorously assessed.

OBJECTIVE
We aim to demonstrate the potentially substantial role of the last age interval on LE variance. We further aim to provide formulae and tools for corrected variance estimation.

METHODS
The delta method is used to derive variance formulae for a range of variance models of the last age interval. Corrected variances are tested on 291 empirical, abridged life tables drawn from Canadian data (2004–2008) spanning provincial, regional, and intra-regional scales.

---

[1] Institut national de santé publique du Québec, 190 blvd Crémazie Est, Montreal, Quebec, Canada. Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Purvis Hall, 1020 Pine Avenue West, Montreal, Quebec, Canada. E-Mail: ernest.lo@mail.mcgill.ca.
[2] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Purvis Hall, 1020 Pine Avenue West, Montreal, Quebec, Canada. E-Mail: dan.vatnik@gmail.com.
[3] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Purvis Hall, 1020 Pine Avenue West, Montreal, Quebec, Canada. E-Mail: andrea.benedetti@mcgill.ca.
[4] Department of Demography, University of Montreal, Pavillon Lionel-Groulx, local C-5014, rue Jean-Brillant, Montreal, Quebec, Canada. E-Mail: robert.bourbeau@umontreal.ca.

RESULTS
The last age interval death count can contribute substantially to the LE variance, leading to overestimates of precision and false positives in statistical tests when using the uncorrected Chiang variance. Overdispersion amplifies the contribution while error in population counts has minimal impact.

CONCLUSIONS
Use of corrected variance formulae is essential for studies that use the Chiang LE. The important role of the last age interval , and hence the life table closure method, on LE variance is demonstrated. These findings extend to other LE-derived metrics such as health expectancy.

CONTRIBUTION
We demonstrate that the last age interval death count can contribute substantially to the LE variance, thus resolving an ambiguity in the scientific literature. We provide heretofore-unavailable formulae for correcting the Chiang LE variance equation.

1. Background

Life expectancy (LE) is a key demographic indicator that provides a summary index of the mortality experience of populations (Shyrock and Siegel 1976). However, it is also a statistical estimator that is computed from underlying random variables: the vital statistics of deaths and births (Brillinger 1986; Chiang 1960) as well as population counts. Each LE estimate thus has a statistical variance that represents a measure of its precision: this precision can be expressed as a standard error, coefficient of variation, or confidence interval, for example. This variance is also essential for statistical tests that are used to ascertain whether or not observed differences in LE are likely caused by random chance. Assessment of LE variance is especially pertinent to smaller, subnational populations[5], where variability in LE can become substantial.

Chiang (Chiang 1960, 1984) first provided a detailed assessment of the statistical properties of life tables and LE. He derived the only analytic equation for the variance of LE, which he expressed, using the delta method, as a weighted sum of the variances of the underlying age-specific survival probabilities. More recent investigations (Eayres and Williams 2004; Scherbov and Ediev 2012; Toson and Baker 2003) have further examined the statistical properties of the Chiang LE estimator down to extremely small population sizes, and established a total population size threshold of 5,000 person-years (PY) above which the LE estimate and its normality remain valid. Due to its

[5] For example, provinces, states, health regions or counties.

accessibility and robustness, the Chiang LE estimator, as well as its associated variance formula, have become an established standard used in both monitoring and scientific studies (von Gaudecker and Scholz 2007; Geruso 2012; Kyte and Wells 2010; Page et al. 2007).

As first noted in 2004 (Eayres and Williams 2004), however, the Chiang variance formula incorrectly omits the variance contribution of the mortality rate of the last age interval[6]. While a subsequent assessment (Toson and Baker 2003) concluded that the magnitude of this contribution to the total LE variance was negligible for populations above 5,000 PY in size, a more recent study (Scherbov and Ediev 2012) suggested a more critical role of the last age interval in the statistical efficiency of LE and the need for further studies. Thus there exists ambiguity in the literature as to the necessity for and impact of a last age interval correction in the Chiang variance formula. Furthermore, none of the above-cited studies have shown or explained how the standard variance formula should be corrected, resulting in further uncertainty for practitioners.

The variance properties of the last age interval have not, in general, been studied, though its contribution to the overall LE variance could be substantial. For example, sparse death counts and the highest mortality rates mostly occur in the last age interval, which lead to elevated statistical variability. Data quality issues in population counts (Bourbeau and Desjardins 2007; Bourbeau and Lebel 2000; Coale and Kisker 1990; Wilmoth and Lundstrom 1996) and heterogeneity or frailty in mortality rates (Bebbington, Lai, and Zitikis 2011; Ting, Yang, and Anderson 2013; Vaupel, Manton, and Stallard 1979) could drive further increases in the variance contribution of the last age interval. There is therefore a general need to quantify and examine the variance contribution of the last age interval under different possible variance models of its corresponding demographic data, which has not previously been done. This essentially entails quantification of the variance of the life table closure method that describes the last age interval contribution to LE.

The objectives of the current study are to assess the variance contribution of the last age interval over a range of variance models, and to evaluate the impact of each model on both the precision and on statistical tests of LE. The delta method used by Chiang will be extended to derive LE variance formulae for each of the variance models considered. Impact assessments will be done on empirical datasets that extend over a range of subnational geographic scales. The error in the Chiang variance formula is first described, and the impact of the omitted last age interval variance contribution is demonstrated. Advanced variance models that account for uncertainty in population counts and mortality rate heterogeneity are next assessed. Finally, the variance of an alternative life table closure method that fits and extrapolates the logistic or 'Kannisto'

---

[6] Silcocks (2001) was the first to note the variance contribution of the last age interval, using a LE estimation method slightly different from that of Chiang

function to the age-specific mortality rates is examined using Monte Carlo simulation; these results are intended to provide an initial assessment of the feasibility of extrapolation closure methods. An Excel tool that computes the life expectancy variances for all models excepting the Kannisto is provided to make the findings of this study more accessible and to further encourage their implementation in practice.

2. Theory

2.1 The delta method approach

A general variance model of LE can be written as $LE = f(y_i | z_j)$, where y_i are identified as independent random variables, and z_j are considered fixed parameters (i.e., measured with negligible error). The LE variance can then be expressed using the delta method (Casella and Berger 2002; ver Hoef 2012) as a weighted sum of the variances of the y_i:

$$\sigma_{LE}^2 \approx \sum_{i=1}^{n} \left(\frac{\partial f}{\partial y_i} \right)^2 \sigma_{y_i}^2,$$

where the weighting factors, $\left(\frac{\partial f}{\partial y_i} \right)^2$, can also be interpreted as the sensitivity of the LE variance to each component variance contribution. The delta method thus provides a useful framework for identifying and quantifying the contributions of each source of stochastic variability, including those of the closure method, to the overall LE variance. This approach will be used to derive formulae for and to quantitatively assess each of the variance models of the last age interval considered in the current study. Key results are outlined below, while Appendix 1 provides more detailed derivations for each variance model. Symbols for life table variables and indices follow the convention used in Chiang (1984), which may differ slightly from standard demographic conventions, in order to facilitate comparison between the standard Chiang variance and the variance formulae of this study. The symbols used to denote the different variances are summarized in Table 1.

Table 1: **Variance models and their associated variance symbols and measures of variance underestimation**

Variance model	Variance model equation	Variance symbol	Measure of variance underestimation
Chiang variance	$f(D_x \mid P_x, D_w, P_w)$	$\sigma^2_{LE.Ch}$	-----
Adjusted Chiang variance	$f(D_x, D_w \mid P_x, P_w)$	$\sigma^2_{LE.adj}$	$\Delta\sigma_{adj.Ch} = \sigma_{LE.adj} - \sigma_{LE.Ch}$
Adjusted Chiang variance with population error	$f(D_x, D_w, P_w \mid P_x)$	$\sigma^2_{LE.adj.pop}$	$\Delta\sigma_{pop.adj} = \sigma_{LE.adj.pop} - \sigma_{LE.adj}$
Adjusted Chiang variance with overdispersion	$f(D_x, D_w \mid P_x, P_w)$	$\sigma^2_{LE.adj.NB}$[7]	$\Delta\sigma_{adj.Ch.NB} = \sigma_{LE.adj.NB} - \sigma_{LE.Ch.NB}$
Kannisto variance	$f(D_x \mid P_x)$	$\sigma^2_{LE.Kn}$	-----

2.2 The 'standard' Chiang variance model

In Chiang's derivation of the LE variance (Chiang 1984), the age-specific death counts are considered the only sources of statistical variability. The variance contribution of the last age interval is completely omitted, however, leading to the following variance model:

$LE_{Ch} = f(D_x \mid P_x, D_w, P_w)$, where D_x and P_x are the age-specific death and population counts corresponding to age group $x = 0 \dots w\text{-}1$, and D_w and P_w are the death and population counts for the last, w^{th}, age interval. The delta method expression for the standard Chiang variance is thus $\sigma^2_{LE.Ch} \approx \sum_{x=0}^{w-1} \left(\frac{\partial LE}{\partial D_x}\right)^2 \sigma^2_{D_x}$. Under the assumption that the death counts are distributed binomially, Chiang derived tractable expressions for the partial derivative weights to obtain the well-known expression for the LE variance:

$$\sigma^2_{LE.Ch} \approx \sum_{x=0}^{w-1} (p_{0x})^2 [(1-a_x)n_x + LE_{x+1}]^2 \left[\frac{q_x^2(1-q_x)}{D_x}\right], \qquad (1)$$

where p_{0x} is the survival probability from age 0 to age x, a_x is the fraction of age interval lived, n_x is the width of the age interval, LE_{x+1} is the life expectancy from age $x + 1$, and q_x is the probability of dying in age interval x. Equation 1 represents the only available analytic formula for the LE variance and has been widely used in studies of LE to estimate its precision and for statistical testing (von Gaudecker and Scholz 2007; Geruso 2012; Kyte and Wells 2010; Stratton et al. 2012). It is evident from this

---

[7] The subscript 'NB' will be used to denote the LE variance with overdispersion; 'NB' refers to the negative binomial distribution that will be used to model overdispersion in the present study. In general, however, overdispersion models other than the negative binomial could also be used.

equation, however, that the Chiang LE variance represents the sum of variance contributions corresponding to each age interval except the last.

2.3 The 'adjusted' Chiang LE variance model

The life table closure method of the Chiang LE estimator assumes that the mortality or hazard rate in the last age interval is constant in time, which results in the following estimate of the person-years contributed: $L_w = \ell_w/M_w = \ell_w P_w/D_w$, where ℓ_w represents the number of survivors at the beginning of the last age interval, and $M_w = D_w/P_w$ represents the mortality rate of the last age interval. As first noted in 2001 (Silcocks, Jenner, and Reza 2001), however, this dependence of LE on D_w entails that its variance contribution should also be accounted for. The corrected variance model for the Chiang LE is then $LE_{adj} = f(D_x, D_w | P_x, P_w)$. Application of the delta method and using a Poisson distribution for D_w yields the following corrected LE variance formula, which has been termed the 'adjusted' Chiang variance in previous studies (Eayres and Williams 2004; Toson and Baker 2003) (cf. Appendix 1 for further details):

$$\sigma^2_{LE.adj} \approx \sum_{x=0}^{w-1} \left(\frac{\partial LE}{\partial D_x}\right)^2 \sigma^2_{D_x} + \left(\frac{\partial LE}{\partial D_w}\right)^2 \sigma^2_{D_w}.$$

The second term in this equation represents the variance contribution of the last age interval death count. Assuming D_w is Poisson distributed such that $\sigma^2_{D_w} = D_w$, and substituting the expression for L_w yields for the correction term (cf. Appendix 1 for further details):

$$\left(\frac{\partial LE}{\partial D_w}\right)^2 \sigma^2_{D_w} = \left[\frac{\partial}{\partial D_w}\left(\frac{L_w}{\ell_0}\right)\right]^2 D_w = \left[\frac{\partial}{\partial D_w}\left(\frac{P_w}{D_w} p_{0w}\right)\right]^2 D_w = \frac{p_{0w}{}^2}{M_w^3 P_w}$$

Therefore,

$$\sigma^2_{LE.adj} \approx \sigma^2_{LE.Ch} + \frac{p_{0w}{}^2}{M_w^3 P_w}. \qquad (2)$$

The following key aspects can be noted: $\sigma^2_{LE.adj}$ is readily estimated by first computing the standard Chiang variance and then adding a correction term, $\sigma^2_{LE.adj} > \sigma^2_{LE.Ch}$, and therefore the standard Chiang variance always underestimates the 'correct' variance, and, the magnitude of the correction term increases with the survival

probability p_{0w} but decreases with increasing mortality rate M_w and population P_w of the last age interval.

2.4 The adjusted Chiang LE variance model with population error

In the statistical analysis of mortality rate metrics, the population or denominator counts are generally assumed to have negligible variability as compared to the death or numerator counts. However, this assumption may be untenable for the last age interval, which comprises the oldest age category of the population, and for which numerous data uncertainty issues have been raised. Identified sources of error include: age exaggeration and underestimation, age heaping, and transcription error (Coale and Kisker 1990). These considerations motivate the following variance model, which includes the variance of population counts of the last age interval: $LE_{adj.pop} = f(D_x, D_w, P_w | P_x)$. The delta method expression for this variance model is as follows:

$$\sigma^2_{LE.adj.pop} \approx \sum_{x=0}^{w-1} \left(\frac{\partial LE}{\partial D_x}\right)^2 \sigma^2_{D_x} + \left(\frac{\partial LE}{\partial D_w}\right)^2 \sigma^2_{D_w} + \left(\frac{\partial LE}{\partial P_w}\right)^2 \sigma^2_{P_w}$$

Detailed assessments of Canadian data using the 'extinct generation' method (Bourbeau and Lebel 2000) indicate a proportional population data error of approximately ±5% in the 90+ age group (cf. Appendix 1, Section 3). This results in a population variance estimate of $\sigma^2_{P_w} \approx \left(\frac{0.05}{2} P_w\right)^2$. Using the estimate for $\sigma^2_{P_w}$ and substituting the expression for L_w yields for this correction term (cf. Appendix 1, Section 3 for further details):

$$\left(\frac{\partial LE}{\partial P_w}\right)^2 \sigma^2_{P_w} = \left[\frac{\partial}{\partial P_w}\left(\frac{P_w}{D_w} p_{0w}\right)\right]^2 \left(\frac{0.05}{2} P_w\right)^2 = \left[\frac{p_{0w}}{D_w}\right]^2 \left(\frac{0.05}{2} P_w\right)^2 = \left(\frac{p_{0w}}{M_w}\right)^2 \left(\frac{0.05}{2}\right)^2$$

Therefore,

$$\sigma^2_{LE.adj.pop} \approx \sigma^2_{LE.adj} + \left(\frac{p_{0w}}{M_w}\right)^2 \left(\frac{0.05}{2}\right)^2 \tag{3}$$

Population error thus contributes an additional term to the adjusted Chiang variance whose magnitude of impact can be readily assessed. Appendix 1, Section 3 provides more details of this derivation.

2.5 The adjusted Chiang LE variance with overdispersion

The *LE* variance model in the presence of overdispersion is identical to that of the adjusted Chiang case, since the sources of stochastic variability remains that of the age-specific death count. Thus $LE_{adj.NB} = f(D_x, D_w | P_x, P_w)$. For the purpose of testing the effect of overdispersion a specific and standard parameterization of the negative binomial (NB) distribution will be used, where age-specific death counts (D_x) are interpreted as the number of events occurring, conditional on a specified probability of death (q_x) and number of survivors at the end of the age interval (denoted P_x^+) (DeGroot 1986; Lloyd-Smith 2007; Manton and Stallard 1981). For $x < w$, the variance of the death count can be readily expressed in terms of the standard binomial variance $\sigma_{D_x}^2$ and the estimated populations at the start (P_x^-) and end (P_x^+) of the interval (cf. Appendix 1, Section 4): $\sigma_{D_x.NB}^2 = D_x \left(\frac{P_x^-}{P_x^+}\right) = \left(\frac{P_x^-}{P_x^+}\right)^2 \sigma_{D_x}^2$. For the last age interval, D_x is interpreted as the number of deaths that occur over an 'observation' period of 5 years (= the number of years of data aggregation). For this interval, $\sigma_{D_w.NB}^2 = D_w \left(\frac{P_w^-}{P_w^+}\right) = \left(\frac{P_w^-}{P_w^+}\right) \sigma_{D_w}^2$ can be written (cf. Appendix 1, Section 4). The *LE* variance in the presence of overdispersion and including the contribution of the last age interval is thus:

$$\sigma_{LE.adj.NB}^2 \approx \sum_{x=0}^{w-1} \left(\frac{\partial LE}{\partial D_x}\right)^2 \sigma_{D_x.NB}^2 + \left(\frac{\partial LE}{\partial D_w}\right)^2 \sigma_{D_w.NB}^2$$

$$= \sum_{x=0}^{w-1} (p_{0x})^2 [(1-a_x)n_x + LE_{x+1}]^2 \frac{q_x^2(1-q_x)}{D_x} \left(\frac{P_x^-}{P_x^+}\right)^2 + \frac{p_{0w}^2}{M_w^3 P_w} \left(\frac{P_w^-}{P_w^+}\right) \quad (4)$$

$$= \sigma_{LE.Ch.NB}^2 + \frac{p_{0w}^2}{M_w^3 P_w} \left(\frac{P_w^-}{P_w^+}\right),$$

where Equations 1 and 2 have been used with the binomial variances replaced by their overdispersed counterparts. As $P_x^- > P_x^+$, Equation 4 shows that overdispersion will inflate each of the component variance contributions to the total LE variance. Also, note that the ratio P_x^-/P_x^+ results in an increasing variance inflation effect with age, consistent with conventional frailty models (Vaupel, Manton, and Stallard 1979).

2.6 The variance model of the Kannisto extrapolation closure method

A detailed empirical study (Thatcher, Kannisto, and Vaupel 1998) has shown that the logistic or 'Kannisto' function accurately describes the mortality rates at the highest ages (80-104 years), overcoming the observed lack of fit of the classical Gompertz function at these ages. Thus the Kannisto extrapolation method fits the logistic function to the age-specific life table mortality rates: $\mu(x) = \propto e^{\beta x}/(1+\propto e^{\beta x})$. The coefficients $\propto, \beta$ are estimated by regression fitting to the age intervals preceding the last, and the function $\mu(x)$ is then extrapolated to estimate the mortality rate profile over the last age interval. The life-years contribution to the LE, which represents the closure method, can then be estimated using standard demographic theory: $L_w = \frac{\ell_w}{\ell_0} \int_{x_w}^{\infty} exp\left(-\int_{x_w}^{x} \mu(y)\, dy\right) dx$.

The variance model for the Kannisto closure method is $LE = f(D_x|P_x)$, $x = 0 \ldots w - 1$, where fitting of data from preceding age intervals has removed any dependence of the LE on data from the last age interval. An analytic expression for $\sigma_{LE.Kn}^2$ is not possible due to the complex dependence of the life years in the last age interval, L_w, on the mortality rates D_x of the preceding age intervals. Therefore in this study the variance of LE using the Kannisto closure method, $\sigma_{LE.Kn}^2$, is estimated using Monte Carlo simulation, with binomial variances used for $\sigma_{D_x}^2$.

3. Data sources and methods

LE variance models were assessed using empirical life table data that spanned three subnational geographic scales: provincial, intra-provincial or regional, and intra-regional; all resulting population strata were above the 5,000 PY threshold in size, as can be seen in Table 2. The provinces and territories of Canada contribute 24[8] 'provincial' and sex-specific strata ranging in size from 78,510–31,880,000 PY (median=2,440,000 PY); life table data were obtained from the Canadian Human Mortality Database (CHMD 2015), aggregated over the years 2004–2008. Stratification over the 18 health administrative regions of Quebec and by sex results in 36 regional strata ranging in size from 26,960–4,794,000 PY (median=807,600 PY). Geographical maps showing the provinces and territories of Canada and the health administrative regions of Quebec are presented in Figures 1a and 1b.

---

[8] Canada comprises 10 provinces and 3 territories; data from the Northwest Territories and Nunavut are aggregated in the Canadian Human Mortality Database to yield life tables corresponding to 12 geographical regions in total.

Table 2: **Total age-aggregated population sizes (in numbers and person-years (PY)) of the three empirical datasets**

		Population size (Numbers and person–years)					
Dataset	# strata	Minimum		Median		Maximum	
Provincial	24	15,700	(78,500 PY)	488,000	(2,440,000 PY)	6,376,000	(31,880,000 PY)
Regional	36	5,344	(26,720 PY)	161,520	(807,600 PY)	958,800	(4,794,000 PY)
Intra-regional	231	1,573	(7,865 PY)	45,220	(226,100 PY)	460,200	(2,301,000 PY)

Further stratification by tertiles defined on the 25th and 75th percentiles of material and social deprivation (Pampalon et al. 2012), including the combined extremes of each dimension, were done resulting in 231 intra-regional strata[9] ranging in size from 7,865–2,301,000 PY (median=226,100 PY). For the regional and intra-regional life tables, death counts were extracted from the Quebec health ministry vital statistics files, while population counts for the middle year of the period were obtained from 2006 census data and adjusted for under-enumeration (Pelletier 2005; Statistics Canada 2006); the first age interval mortality rate is estimated by the ratio of deaths to the number of live births (Brown 1993). Life tables were calculated using the Chiang method (Chiang 1984); in the implementation of the Kannisto closure method, the Chiang method was used to estimate the life years contributed for all age intervals except the last. Five-year age intervals were used except for the first (0–1) and last (90+) intervals. For the first age interval, the fraction of interval lived was set to 0.1. All life table and subsequent statistical analyses were done using R (version 2.15) statistical software.

---

[9] Regions 10, 17, and 18 are excluded here due to their smaller population size, as well as 9 strata for which population counts of zero occurred in particular age groups.

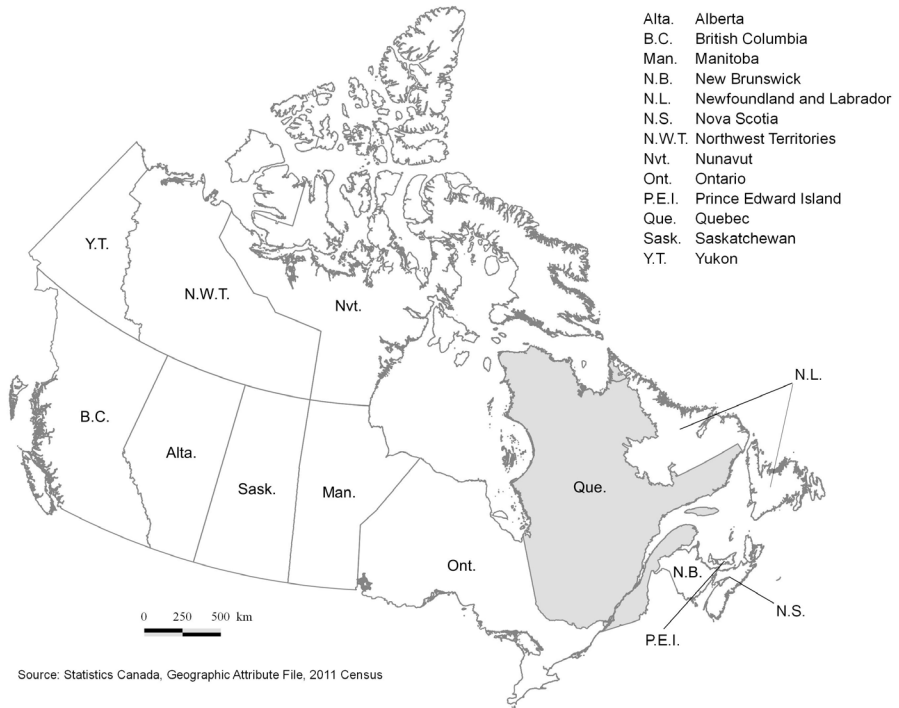**Figure 1a):   Provinces and territories of Canada (Quebec shaded in gray)**

Alta.   Alberta
B.C.   British Columbia
Man.   Manitoba
N.B.   New Brunswick
N.L.   Newfoundland and Labrador
N.S.   Nova Scotia
N.W.T. Northwest Territories
Nvt.   Nunavut
Ont.   Ontario
P.E.I. Prince Edward Island
Que.   Quebec
Sask.   Saskatchewan
Y.T.   Yukon

Source: Statistics Canada, Geographic Attribute File, 2011 Census

Figure 1b):   The 18 health administrative regions of Quebec

01 - Bas-Saint-Laurent
02 - Saguenay-Lac-Saint-Jean
03 - Capitale-Nationale
04 - Mauricie et Centre-du-Québec
05 - Estrie
06 - Montréal
07 - Outaouais
08 - Abitibi-Témiscamingue
09 - Côte-Nord
10 - Nord-du-Québec
11 - Gaspésie-Iles-de-la-Madeleine
12 - Chaudière-Appalaches
13 - Laval
14 - Lanaudière
15 - Laurentides
16 - Montérégie
17 - Nunavik
18 - Terres-Cries-de-la-Baie-James

Source: Minister of Health and Social Services of Quebec, Health and social limits, 2015

The difference in LE standard error with and without a particular variance contribution of the last age interval is termed $\Delta\sigma$ $(= \sigma_{corrected} - \sigma_{uncorrected})$ and represents the resulting degree of underestimation of the variance of LE; $\Delta\sigma$ is also directly proportional to the magnitude of under-coverage of the associated confidence intervals. Table 1 lists the equation for $\Delta\sigma$, for each variance model, when applicable. Scatterplots of $\Delta\sigma$ vs. population size are used to examine the distribution of error or variance underestimation corresponding to each variance model, and to examine possible scaling trends and classification thresholds. 1–1 scatterplots of the 'corrected' vs. 'uncorrected' LE standard errors are used to further examine the magnitude of each

variance correction relative to the absolute magnitude of the LE standard error, and to provide comparison of corrected vs. uncorrected standard error magnitudes.

The global impact of the last age interval variance corrections is estimated using the prevalence at which elevated $\Delta\sigma$ occurs over the 3 empirical datasets using two thresholds: $\Delta\sigma > 0.02$ and $\Delta\sigma > 0.1$. The former threshold was established by Toson and Baker (Toson and Baker 2003) and will be critically re-evaluated in the current study; the latter threshold represents a more liberal value that corresponds to a substantial increase in LE confidence interval coverage of approximately 0.4 years.

The false positive rate (or type 1 error rate) (Gravetter and Wallnau 2002) measures the increase in the number of erroneous statistical test results that occur due to the omission of a last age interval variance component (and concomitant underestimation of the standard error of the LE estimator). It equals the ratio of the number of false positives to the total number of statistical tests performed, when an 'uncorrected' variance formula is used relative to a 'corrected' one. False positive rates are estimated using statistical tests done in each of the three datasets, which comprise: tests of LE differences relative to the Canadian average in provincial-level strata, tests of LE differences relative to the Quebec average in regional strata, and tests of LE inequalities via comparison of LE between extremes (i.e., high vs. low) of material, social, and combined deprivation tertiles in intra-regional strata. Two significance levels or nominal false positive rates, $\alpha = 0.05$ (or 5%) and $\alpha = 0.01$ (or 1%), are considered: the former represents a standard significance level, while the latter represents a frequently used, more conservative level.

4. Results

4.1 The adjusted Chiang variance model

Figure 2a shows $\Delta\sigma_{adj.Ch} = \sigma_{LE.adj} - \sigma_{LE.Ch}$ plotted vs. population size for the adjusted Chiang variance model. Crosses, triangles, and gray circles represent results for the provincial, regional, and intra-regional strata respectively, while the dashed lines indicate the 0.02 and 0.1 threshold values. The prevalence of $\Delta\sigma_{adj.Ch}$ exceeding these two thresholds is also presented in Table 3a.

Overall, it can be seen that the last age interval can contribute substantially to the LE variance. Approximately 49% of strata exhibit $\Delta\sigma_{adj.Ch}$ higher than the 0.02 upper limit (cf. Table 3a) reported in previous studies by Toson and Baker, despite all strata exceeding 5,000 PY in size. $\Delta\sigma_{adj.Ch}$ frequently attains values greater than 0.1 years, with several extreme values found near 7 years, as reflected by a prevalence of $\approx$9% of $\Delta\sigma_{adj.Ch} > 0.1$ (cf. Table 3a).

Figure 2a shows an increasing trend for $\Delta\sigma_{adj.Ch}$ with decreasing population size, though considerable variability in the results prevents establishment of a clear threshold that would reliably classify strata into low vs. high $\Delta\sigma_{adj.Ch}$, especially over regional and sub-regional strata. This trend with population size is also evidenced in Table 3a, which shows that the prevalence of elevated $\Delta\sigma_{adj.Ch}$ increases with decreasing geographic scale, from 4.2% to 58.4% corresponding to $\Delta\sigma_{adj.Ch} > 0.02$, and from 0.0% to 10.4% corresponding to $\Delta\sigma_{adj.Ch} > 0.1$. $\Delta\sigma_{adj.Ch}$ also exhibited general scaling trends when plotted against a range of other key life table parameters, though no clear classification thresholds could be identified here either (cf. Appendix 2).[10]

Figure 2b shows the scatterplot of the standard Chiang vs. adjusted Chiang standard errors. The adjusted Chiang standard error varies greatly in size, from ≈0.03 to 11 years, reflecting the range of population sizes and LE magnitudes spanned by the empirical datasets. The magnitude of the last age interval variance contribution is represented by the rightward shift in data points from the 1−1 (dotted) line. It can be seen that for many strata, especially from the intra-regional dataset, there is a sizeable last age interval variance contribution relative to the magnitude of the LE standard error. This contribution tends to increase with increasing SE with notably large contributions of $\Delta\sigma_{adj.Ch}$ ≈1−10 years occurring in the LE estimates with the largest standard errors.

Table 3b lists the false positive rates that result from use of the standard Chiang variance relative to the adjusted Chiang variance, for the statistical tests previously described, over the 3 datasets. Use of the standard Chiang variance results in substantial overall increases in false positive rates of 4.9% and 4.2% over nominal false positive rates of α = 0.05 (5%) and α = 0.01 (1%), respectively. No false positives occurred in statistical tests done in the province-level strata, while the rates for the region-level strata were 5.6% and 5.6% for α = 0.05 and 0.01, and those for the intra-regional strata were 4.8% and 6.0% respectively.

---

[10] Note that due to the tight correspondence between population size and LE standard error, scaling trends with respect to population size are quite similar to those with respect to the LE standard error (cf. Appendix 5).

Figure 2: **(a)** $\Delta\sigma_{adj.Ch}$ **vs. population size and (b)** $\sigma_{LE.Ch}$ **vs.** $\sigma_{LE.adj}$
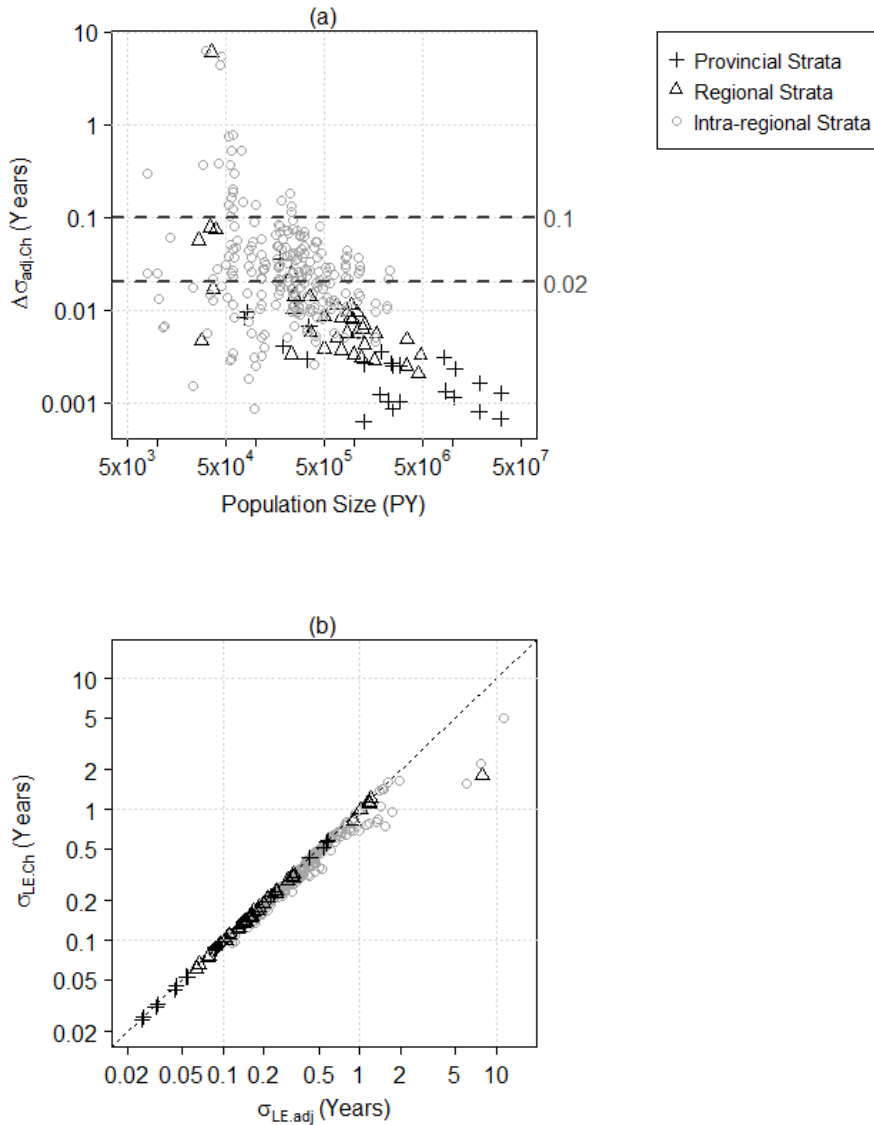
Table 3a): **Proportion of strata with elevated $\Delta\sigma_{adj.Ch}$**

Dataset	# of Strata	% $\Delta\sigma_{adj.Ch}$ > 0.02	% $\Delta\sigma_{adj.Ch}$ > 0.1
Provincial	24	4.2% (1/24)	0.0% (0/24)
Regional	36	13.9% (5/36)	2.8% (1/36)
Intra-regional	231	58.4% (135/231)	10.4% (24/231)
Total	291	48.5% (141/291)	8.6% (25/291)

Table 3b): **Induced false positive rates for $\sigma^2_{LE.Ch}$ with respect to $\sigma^2_{LE.adj}$**

Dataset	# of Statistical Comparisons	False positive rate, α=0.05	False positive rate, α=0.01
Provincial	24	0.0% (0/24)	0.0% (0/24)
Regional	36	5.6% (2/36)	5.6% (2/36)
Intra-regional	83	4.8% (4/83)	6.0% (5/83)

4.2 The adjusted Chiang variance model with population error

$\Delta\sigma_{pop.adj} = \sigma_{LE.adj.pop} - \sigma_{LE.adj}$, shown plotted in Figure 3a, represents the incremental variance contribution of population error in the last age interval, relative to the adjusted Chiang variance, estimated over the 3 datasets. In contrast to $\Delta\sigma_{adj.Ch}$, $\Delta\sigma_{pop.adj}$ increases with population size due to the proportional dependence of the population variance on P_w as indicated by the last term in Equation 3. Overall, however, the magnitude of $\Delta\sigma_{pop.adj}$ is relatively small, reaching an upper limit of ≈0.02 years only when P_{ttl} exceeds 1.5×10^6 PY. As shown in Table 4a, the prevalence of $\Delta\sigma_{pop.adj}$>0.02 is 1.7%, indicating that for a relatively small number of cases, only a slight increase in LE standard errors over that of $\Delta\sigma_{adj.Ch}$ occurs, with the largest strata from the provincial and intra-regional datasets contributing these additional cases. Inclusion of population error results in no cases with $\Delta\sigma_{pop.adj}$>0.1.

Figure 3b shows the adjusted Chiang standard error ($\sigma_{LE.adj}$) plotted against the adjusted Chiang standard error with population error included ($\sigma_{LE.adj.pop}$). It can be seen that the incremental contribution of population error is negligible except in the case of the largest population strata ($P_{ttl} \gtrsim 1.5\times10^6$ PY) where the pre-existing small LE standard errors are increased by proportionately large amounts. Inclusion of

population error results in a slight incremental impact on statistical tests, with a 1.4% increase in false positive rates relative to the adjusted Chiang variance model, as shown in Table 4b.

Figure 3: (a) $\Delta\sigma_{pop.adj}$ vs. population size and (b) $\sigma_{LE.adj}$ vs. $\sigma_{LE.adj.pop}$
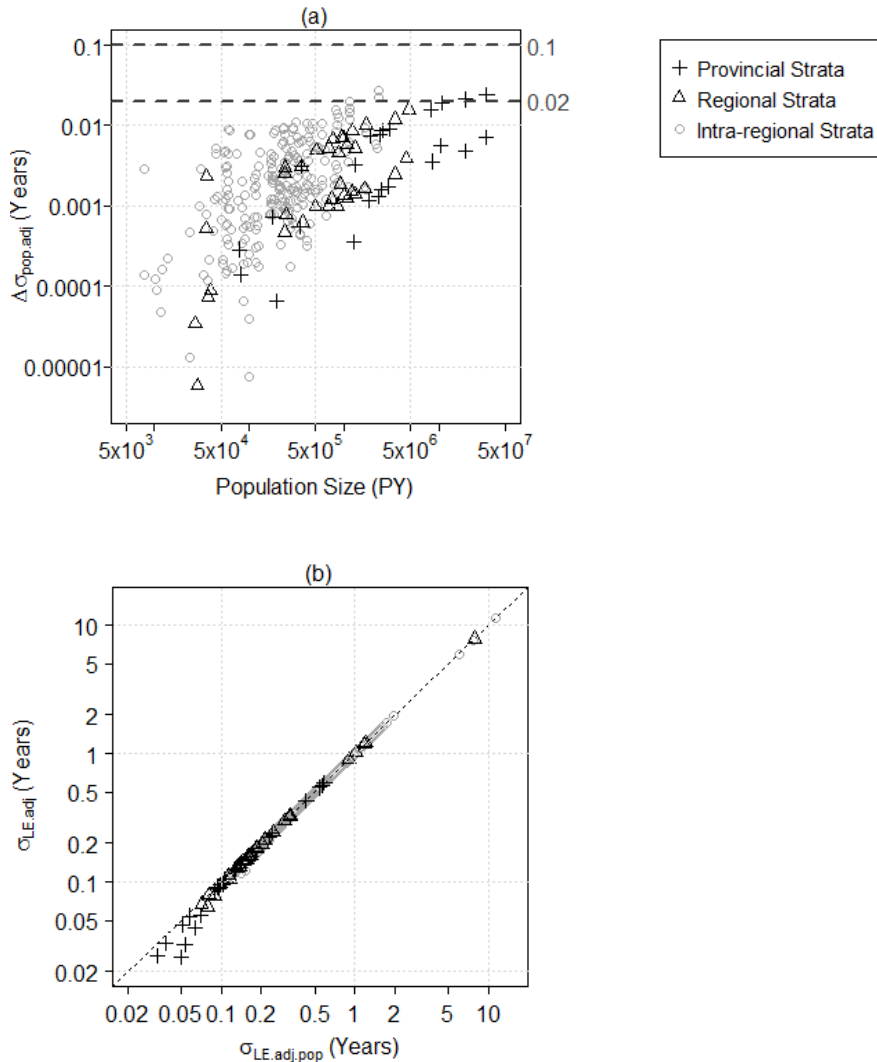
Table 4a): **Proportion of strata with elevated $\Delta\sigma_{pop.adj}$**

Dataset	# of Strata	% $\Delta\sigma_{pop.adj}$ > 0.02	% $\Delta\sigma_{pop.adj}$ > 0.1
Provincial	24	8.3% (2/24)	0.0% (0/24)
Regional	36	0.0% (0/36)	0.0% (0/36)
Intra-regional	231	1.3% (3/231)	0.0% (0/231)
Total	291	1.7% (5/291)	0.0% (0/291)

Table 4b): **Induced false positive rates for $\sigma^2_{LE.adj}$ with respect to $\sigma^2_{LE.adj.pop}$**

Dataset	# of Statistical Comparisons	False positive rate, α=0.05	False positive rate, α=0.01
Provincial	24	8.3% (2/22)	0.0% (0/20)
Regional	36	0.0% (0/26)	2.8% (1/24)
Intra-regional	83	0.0% (0/53)	1.2% (1/46)
Total	143	1.4% (2/101)	1.4% (2/90)

4.3 The last age interval variance contribution in the presence of overdispersion

Figure 4a shows the variance contribution of the last age interval in the presence of overdispersion ($\Delta\sigma_{adj.Ch.NB} = \sigma_{LE.adj.NB} - \sigma_{LE.Ch.NB}$) plotted vs. population size. The magnitude of $\Delta\sigma_{adj.Ch.NB}$ is in general higher than that of its non-overdispersed counterpart $\Delta\sigma_{adj.Ch}$, due to the P_w^-/P_w^+ factor contributed mathematically by overdispersion in the last age interval. This is reflected in the markedly higher overall prevalences of 80.1% of $\Delta\sigma_{adj.Ch.NB}$ > 0.02 and 19.9% of $\Delta\sigma_{adj.Ch.NB}$ > 0.1 (cf. Table 5a).

$\Delta\sigma_{adj.Ch.NB}$ decreases with increasing population size, following trends that resemble closely those of $\Delta\sigma_{adj.Ch}$. The prevalence of elevated $\Delta\sigma_{adj.Ch.NB}$ also increases throughout the three datasets with decreasing population size, from 16.7% to 91.3% for $\Delta\sigma_{adj.Ch.NB}$ > 0.02, and from 8.3% to 22.5% for $\Delta\sigma_{adj.Ch.NB}$ > 0.1 (cf. Table 5a).

The scatterplot of $\sigma_{LE.adj.NB}$ vs. $\sigma_{LE.Ch.NB}$ in Figure 4b illustrates that the size of the LE standard errors themselves ranges from ≈0.03 to 16 years over the 3 datasets and

has thus been amplified by overdispersion. $\Delta\sigma_{adj.Ch.NB}$ is also proportionately larger overall, as shown by the more pronounced rightward shift in points from the dotted black (1–1) line.

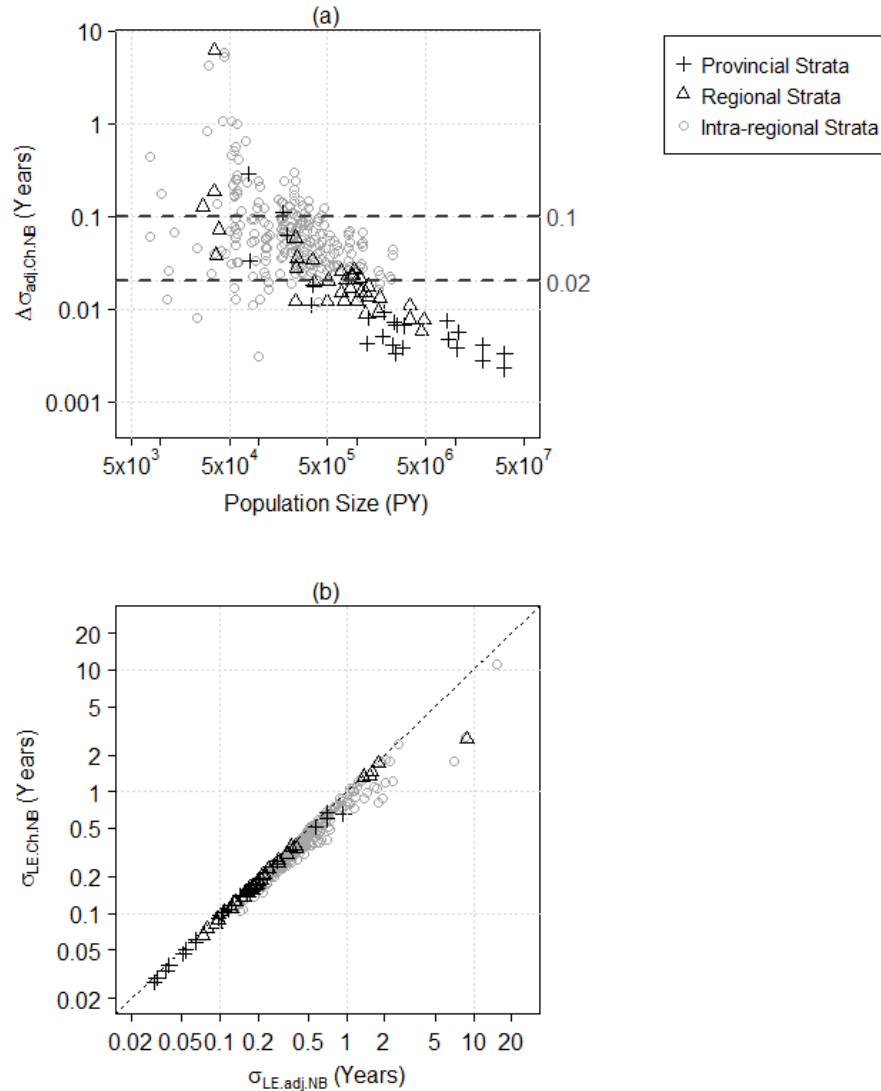**Figure 4:** **(a) $\Delta\sigma_{adj.Ch.NB}$ vs. population size and (b) $\sigma_{LE.Ch.NB}$ vs. $\sigma_{LE.adj.NB}$**

Table 5b (in comparison with Table 3b) shows that the impact of the last age interval variance on statistical tests is further amplified in the presence of overdispersion. This is evidenced by substantially higher increases in false positive rates by 12.6% and 8.4% with respect to nominal false positive rates or significance levels of 0.05 (5%) and 0.01 (1%), respectively. This overall increase in false positive rates is produced by corresponding increases in virtually all dataset and α combinations, which show increases in false positive rates ranging up to 14.5%.

Table 5a):   **Proportion of strata with elevated $\Delta\sigma_{adj.Ch.NB}$**

Dataset	# of Strata	% $\Delta\sigma_{adj.Ch.NB}$ > 0.02	% $\Delta\sigma_{adj.Ch.NB}$ > 0.1
Provincial	24	16.7% (4/24)	8.3% (2/24)
Regional	36	50.0% (18/36)	11.1% (4/36)
Intra-regional	231	91.3% (211/231)	22.5% (52/231)
Total	291	80.1% (233/291)	19.9% (58/291)

**Table 5b):**   **Induced false positive rates for $\sigma^2_{LE.adj.NB}$ with respect to $\sigma^2_{LE.Ch.NB}$**

Dataset	# of Statistical Comparisons	False positive rate, α=0.05	False positive rate, α=0.01
Provincial	24	8.3% (2/22)	0.0% (0/0)
Regional	36	11.1% (4/27)	13.9% (5/22)
Intra-regional	83	14.5% (12/55)	8.4% (7/45)
Total	143	12.6% (18/104)	8.4% (12/87)

4.4 Comparative performance of the Kannisto extrapolation closure method

Closure of the life table with the Kannisto extrapolation procedure results in an LE estimator that is no longer equivalent to the Chiang LE. Thus the resulting Kannisto LE variance cannot be represented as an additive variance contribution to the standard Chiang variance, as was done for the other variance models tested. However, the performance of the Kannisto LE variance can be roughly assessed through scatterplots of the Kannisto LE standard error, $\sigma_{LE.Kn}$, vs. the adjusted Chiang standard error, $\sigma_{LE.adj}$. Figure 5a shows $\sigma_{LE.Kn}$ vs. $\sigma_{LE.adj}$ for the subset of cases in which $\Delta\sigma_{adj.Ch}$>0.1 and in which there is an elevated variance contribution of the last age interval. It can be seen that for these cases the Kannisto method reduces or stabilizes the

variance contribution of the last age interval, as indicated by the overall rightward shift in points from the dashed (1−1) line in the Figure. In particular, substantial reductions in standard error of up to ≈10 years occur when $\sigma_{LE.adj}$ becomes large (>1). Figure 5b shows that for $\Delta\sigma_{adj.Ch} < 0.1$, $\sigma_{LE.Kn}$ is comparable to $\sigma_{LE.adj}$, which is expected since the variance contribution of the last age interval is already small.

Figure 5: $\boldsymbol{\sigma_{LE.Kn}}$ vs. $\boldsymbol{\sigma_{LE.adj}}$ for (a) $\Delta\boldsymbol{\sigma_{adj.Ch}} \geq \boldsymbol{0.1}$ and (b) $\Delta\boldsymbol{\sigma_{adj.Ch}} < \boldsymbol{0.1}$[11]

---
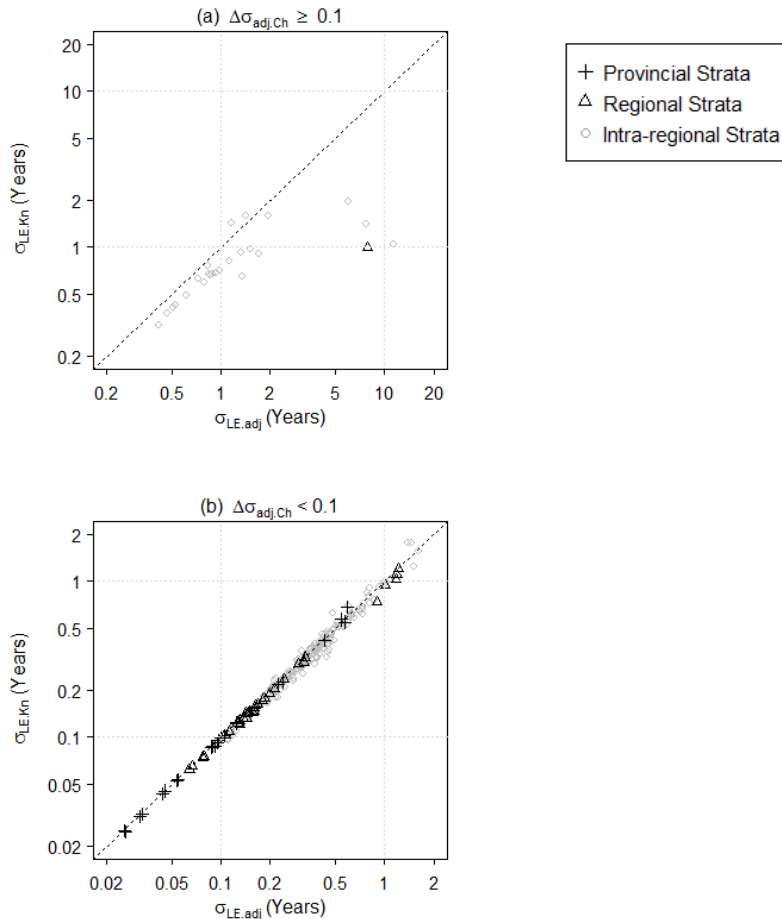
[11] Results have been 'split' by level of $\Delta\sigma_{adj.Ch}$ so that deviations of data points from the 1−1 line are more clearly and readily discerned.

5. Discussion

The current study demonstrates that the last age interval can contribute substantially to the LE variance when using the Chiang LE estimator. It is firstly shown that inclusion of the variance contribution of the last age interval mortality count can lead to increases in the LE standard error that greatly exceed the previously reported 0.02 upper limit (Toson and Baker 2003), even for populations substantially larger than 5,000 PY. It is further demonstrated that overestimation of the precision of LE is widespread and leads to the substantial inflation of false discovery rates or type 1 error in statistical comparisons at subnational scales.

Assessment of an extended variance model indicates that population error in the last age interval contributes negligibly to the LE variance for population sizes below 1.5 x 10^6 PY. For populations > 1.5 x 10^6 PY, population error can make a proportionately substantial contribution to the total variance, although the absolute variance contribution remains small (< 0.03 years). In light of these results and of the likely improvements in population data quality since 1971–1991 (from which conservative estimates of a 5% proportional error in population counts were derived), it is likely that the population error of the last age interval can be neglected for purposes of LE variance estimation. Despite representing a known and major source of error, assessment of the last age interval population count error on LE variance has not previously been done. Our study addresses this gap in knowledge and provides a clear recommendation for practitioners.

The variance contribution of the last age interval increases in the presence of overdispersion, leading to marked increases in both the prevalence of severe variance underestimation and inflated false positive rates when the uncorrected Chiang variance is used. Overdispersion, also known as heterogeneity or frailty, is likely present to some degree in most mortality data, and manifests in the observed mortality trends at the highest ages (Andreev and Bourbeau 2006; Bebbington, Lai, and Zitikis 2011; Ting, Yang, and Anderson 2013). Thus the overdispersed LE variance results of this study may represent a more realistic scenario, which further underscores the substantial impact of the last age interval on variance.

Monte Carlo simulations show that the Kannisto extrapolation closure method reduces or stabilizes the variance contribution of the last age interval. However, an analytic representation of the corresponding LE variance is no longer possible due to the increased complexity of the last age interval life-years expression. These results suggest that while extrapolation closure methods may improve the precision of LE, algorithms to simplify their variance calculation for practical use would be needed. Further examination and development of alternative closure methods in general may represent an important area of future research.

Inclusion of the variance contribution of the last age interval death count is thus found to be essential when using the Chiang LE estimator at subnational scales. The substantial overestimation of precision and false positives that results from use of the uncorrected Chiang LE variance can lead to errors that have both scientific and policy consequences. As reliable thresholds to classify population strata in regard to the expected magnitude of the last age interval variance are not possible, the mathematical correction of each LE variance estimate is recommended. To this end, the simple, additive formulae provided in this manuscript can be used by practitioners to readily correct existing or planned variance estimates. The Excel tool provided in this study demonstrates the implementation of each of these corrections on a sample life table, and is intended to further facilitate implementation.

To summarize the findings of our study for operational purposes, the adjusted Chiang variance (Equation 2, Section 2.3, $\sigma^2_{LE.adj}$) should be used in place of the standard Chiang formula for estimating the LE variance. The adjusted Chiang variance with overdispersion (Equation 4, Section 2.5, $\sigma^2_{LE.adj.NB}$) is recommended when the choice has already been made to model overdispersion in mortality counts in the other age intervals. The addition of population error (Section 2.3) is found to affect the LE variance negligibly and need not be used. The variance model of the Kannisto LE (Section 2.6) is valid but requires Monte Carlo simulation to evaluate, and applies only when the Kannisto life table closure method is used in place of the Chiang closure method.

The findings of this study extend beyond the basic life expectancy from birth (LE) metric that was examined as a test case. The derived variance corrections are readily adapted (cf. Appendix 1) for the estimation of LE from advanced ages (i.e., 25 or 65 years), for example, and for which the last age interval contribution will be proportionately even higher. The Chiang method also forms the basis for a broad spectrum of health expectancy (HE) metrics that have gained global acceptance for the assessment and monitoring of population health (Robine, Romieu, and Cambois 1999; Salomon et al. 2012) and which are frequently presented together with LE estimates (Bajekal 2005; Geronimus et al. 2001; Salomon et al. 2012). It is recommended that correction terms due to the variance contribution of the last age interval be applied as needed to these metrics as well, to avoid the same substantial variance underestimation problems. An example of this is shown in Appendix 3 for health adjusted life expectancy (HALE) estimated at the regional scale.

5.1 Study limitations

Although tests were done on a wide range of empirical data, these data are not completely general and do not represent the full range of possible life tables. Nevertheless, the results are sufficient to demonstrate the substantial impact of the last age interval on LE variance, which is the principal objective of the study. Overall, the great variability in stratum-specific results observed in the current study, and the demonstrated inadequacy of the previously identified 0.02 limit (which was estimated from simulations of a single test population), underscore the need for empirical datasets when assessing the performance of life-table-based metrics.

Life table data were restricted to those of Canada (for province level analyses) and Quebec (for regional and intra-regional analyses), though these data are expected to be comparable to other developed nations. LE in fact varies quite widely (from 64.1 to 94.5 years) over strata, indicating a broad range of demographic conditions, some of which may resemble those of other countries at different stages of development. In particular, Figure A-2.1e illustrates that substantial variance underestimation may occur regardless of the magnitude of estimated LE, and thus indicates the likely importance of the last age interval for non-Canadian datasets.

In reality, the true or correct LE variance is inherently unknown, since LE can only be sampled or observed once for a given time period and population. Therefore the increase in false positive rate was estimated approximately by comparing statistical tests using 'uncorrected' vs. 'corrected' variances. This likely results in conservative estimates of false positive rate increase and thus of impact, since 'corrected' variances are likely underestimates of the true LE variance due to the presence of additional, unaccounted-for sources of variability. The hypothetical use of true LE variances would therefore tend to further increase the impact of the estimated LE variance corrections and thus further strengthen the conclusions the study. Overall, the 'corrected' LE variance formulae derived in this study account for the stochastic contribution of age-specific death counts according to established statistical theory (Brillinger 1986; Chiang 1960; Keyfitz 1966), and thus are expected to be reasonable estimates of the true LE variance.

Variability in the last age interval population counts (σ_{P_w}) is likely produced by age misreporting and other sources of error, whose exact mechanisms and thus probability distributions are in fact largely unknown. Nevertheless, the finding that the impact of σ_{P_w} on the overall LE variance is likely negligible is both general and robust with respect to assumptions made in its estimation. Firstly, the delta method applies to the variance of LE and its component variables and is independent of the actual underlying probability distributions. From the delta method it can be shown mathematically that the LE variance is much less sensitive to population error as

compared with error in death counts. Using an empirically based 5% proportional error, it can be further demonstrated that the variance contribution of population counts becomes comparable to that of death counts only for very large ($\approx 1.2 \times 10^6$) population sizes, where the total LE variance is likely small. Finally, additional validation analyses have shown that the impact of population error remains negligible despite varying the assumed coverage associated with the observed empirical error, as well as consideration of corresponding 'opposing' population count error in younger age intervals. Detailed discussion and analyses of these issues are presented in Appendix 3.

The delta method utilizes a first order Taylor expansion of the function for which the variance is estimated. For this reason, the various derived expressions for variance are in fact first order approximations. However, validation studies that have compared the delta method standard error with Monte Carlo simulated standard errors have shown that the delta method is accurate (Eayres and Williams 2004). The delta method was used by Chiang in his original derivation of the LE variance (though the contribution the last age interval was neglected): it also sees widespread use for variance estimation in biostatistics in general (Cox 2005).

Performance of the Chiang LE variance was examined using a closure age of 90 years, which is the established norm in the Quebec Public Health Network (MSSS 2011) and ensures robust LE estimation over datasets that include small strata. At provincial or national scales, however, sufficiently large age-specific population sizes and data quality often permit LE estimation using elevated closure ages of up to 100 years or higher (CHMD 2015; Martel et al. 2012; Wilmoth et al. 2007). An increase in closure age to 100 years would result in a greatly reduced life table survival probability from birth to the last age interval (p_{0w}), which would in turn reduce the magnitude of the last age interval contribution to the LE variance (cf. Equation 2). Furthermore, due to the inverse scaling with population size, the absolute magnitudes of the total LE variance and the last age interval contribution become quite small. Thus it must be emphasized that the scope of the current study pertains to LE estimation at subnational scales (down to population sizes of 5,000 PY), where the LE variance is relatively substantial and closure ages of 90 years or less are typically used.

Conversely, reduction of the life table age to below 90 years could also represent a possible strategy to alleviate large variance contributions from the last age interval, by effectively increasing the last age interval population size and thereby reducing the frequency of sparse death counts. Validation tests show, however, that the last age interval variance contribution ($\Delta\sigma_{Ch.adj}$) remains substantial and widespread even when the life table closure age is reduced to 85 years (cf. Appendix 4). A further test demonstrates the reduction in the variance contribution in the last age interval when the life table closure age is increased to 95 years (cf. Figure A-4.2, Appendix 4); there is a

substantial increase in excluded strata due to zero age-specific population counts or zero death counts in the last age interval for this case however[12].

While a specific though widely used parameterization of the negative binomial distribution was used to describe heterogeneity in death counts, the actual degree of heterogeneity in real populations is unknown and difficult to estimate. However, the negative binomial parameterization used in the current study results in a variance inflation effect that increases with age, consistent with the expected increases in frailty in mortality rates due to survivorship effects (Vaupel, Manton, and Stallard 1979), and is thus a reasonable, representative model for heterogeneity. As a result, the general finding of the study – that the impact of the last age interval on the LE variance is amplified in the presence of overdispersion – is expected to hold for other realistic heterogeneity or frailty models.

Finally, the present study focuses on the variance properties of the Chiang LE estimator, which is but one of many existing LE estimators. It should be noted, however, that the Chiang closure method is identical to that used by numerous other classic LE methods including the 'actuarial', Greville, Sirken, Keyfitz-Frauenthal, and Schoen approaches (Greville 1943; Keyfitz and Frauenthal 1975; Schoen 1978; Siegel 2012). Thus the findings regarding the last age interval variance are applicable to a range of LE estimators. More generally, the Chiang method remains the most widely used and accessible procedure for the estimation of LE and its variance, and thus represents an important subject of study.

6. Conclusions

The present study represents the first detailed assessment of the variance properties of the last age interval and its impact on LE variance. The delta method framework has been extended to model population data error and mortality rate overdispersion properties of the underlying vital rates. Tests done on a range of empirical sub-national life table data demonstrate that the variance contribution of the last age interval can lead to severe underestimation of the total LE variance, as well as substantial increases in the false positive rate in statistical tests. It is hoped that this study will thus unambiguously demonstrate to practitioners and users the need to correct the standard Chiang LE variance. To facilitate implementation, accessible, additive correction formulae are provided for correcting the standard Chiang LE variance. Additionally, an Excel tool is provided that computes the life expectancy variances for all variance models (excepting the Kannisto) for a sample life table. The Chiang LE estimator is a widely accepted

---

[12] Although methods (such as imputation) exist for dealing with these strata, examination of this issue is beyond the scope of the present study.

standard for use by demographers, epidemiologists, and public health officers, due to its accessibility, ease of computation, robustness for smaller populations ($\gtrsim$ 5,000 PY), and available variance formula. It is hoped that the present study will contribute toward the improved validity of the numerous previous, ongoing, and future analyses done using this estimator.

7. Acknowledgements

References

Andreev, K.F. and Bourbeau, R. (2006). Frailty modeling of Canadian and Swedish mortality at adult and advanced ages. Unpublished manuscript.

Bajekal, M. (2005). Healthy life expectancy by area deprivation: Magnitude and trends in England, 1994–1999. *Health Statistics Quarterly* 25: 18–27.

Bebbington, M., Lai, C.D., and Zitikis, R. (2011). Modelling deceleration in senescent mortality. *Mathematical Population Studies* 18(1): 18–37. doi:10.1080/0889848 0.2011.540173.

Bourbeau, R. and Desjardins, B. (2007). Mortality at extreme ages and data quality: The Canadian experience. In: Robine, J.-M., Crimmins, E.M., Horiuchi, S., Zeng, Y. (eds.). *Human longevity, individual life duration, and growth of the oldest-old population*. Berlin: Springer: 167–185. doi:10.1007/978-1-4020-4848-7_8.

Bourbeau, R. and Lebel, A. (2000). Mortality statistics for the oldest-old: An evaluation of Canadian data. *Demographic Research* 2(2): doi:10.4054/demres.2000.2.2.

Brillinger, D.R. (1986). The natural variability of vital rates and associated statistics. *Biometrics* 42(4): 693–734. doi:10.2307/2530689.

Brown, R.L. (1993). *Introduction to the mathematics of demography*. New Hartford: ACTEX Publications.

Casella, G. and Berger, R.L. (2002). *Statistical inference*. Pacific Grove: Duxbury.

Chiang, C.L. (1960). A stochastic study of the life table and its applications: II. Sample variance of the observed expectation of life and other biometric functions. *Human Biology* 32(3): 221–238.

Chiang, C.L. (1984). *The life table and its applications*. Malabar: Robert E. Krieger.

CHMD (2015). Canadian Human Mortality Database [electronic resource]. http://www.bdlc.umontreal.ca/chmd/.

Coale, A.J. and Kisker, E.E. (1990). Defects in data on old-age mortality in the United States: New procedures for calculating mortality schedules and life tables at the highest ages. *Asian and Pacific Population Forum* 4(1): 1–32.

Cox, C. (2005). Delta method. In: Armitage, P. and Colton, T. (eds.). *Encyclopedia of biostatistics*. New York: John Wiley: 1540–1542. doi:10.1002/0470011815. b2a15029.

DeGroot, M.H. (1986). *Probability and statistics*. Reading: Addison-Wesley.

Eayres, D. and Williams, E.S. (2004). Evaluation of methodologies for small area life expectancy estimation. *Journal of Epidemiology and Community Health* 58: 243–249. doi:10.1136/jech.2003.009654.

Geronimus, A.T., Bound, J., Waidmann, T.A., Colen, C.G., and Steffick, D. (2001). Inequality in life expectancy, functional status, and active life expectancy across selected black and white populations in the United States. *Demography* 38(2): 227–251.

Geruso, M. (2012). Black-White disparities in life expectancy: How much can the standard SES variables explain? *Demography* 49(2): 553–574. doi:10.1007/s13 524-011-0089-1.

Gravetter, F.J. and Wallnau, L.B. (2002). *Essentials of statistics for the behavioral sciences*. Pacific Grove: Wadsworth.

Greville, T.N.E. (1943). Short methods of constructing abridged life tables. *Record of the American Institute of Actuaries* 32(1): 29–42.

Horsman, J., Furlong, W., Feeny, D., and Torrance, G. (2003). The Health Utilities Index (HUI): Concepts, measurement properties and applications. *Health and Quality of Life Outcomes* 1(54). doi:10.1186/1477-7525-1-54.

Keyfitz, N. (1966). Sampling variance of standardized mortality rates. *Human Biology* 38(3): 309–317.

Keyfitz, N. and Frauenthal, J. (1975). An improved life table method. *Biometrics* 31(4): 889–899. doi:10.2307/2529814.

Kyte, L. and Wells, C. (2010). Variations in life expectancy between rural and urban areas of England, 2001–07. *Health Statistics Quarterly* 46: 25–50.

Lloyd-Smith, J.O. (2007). Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. *PLoS One* 2(2): e180. doi:10.1371/journal.pone.0000180.

Loukine, L., Waters, C., Choi, B.C., and Ellison, J. (2012). Impact of diabetes mellitus on life expectancy and health-adjusted life expectancy in Canada. *Population Health Metrics* 10(1): 7.

Manton, K.G. and Stallard, E. (1981). Methods for the analysis of mortality risks across heterogeneous small populations: Examination of space-time gradients in cancer mortality in North Carolina counties 1970–75. *Demography.* 18(2): 217–230.

Manuel, D.G., Schultz, S.E., and Kopec, J.A. (2002). Measuring the health burden of chronic disease and injury using health adjusted life expectancy and the Health Utilities Index. *Journal of Epidemiology and Community Health* 56(11): 843–850.

Martel, L., Provost, M., Lebel, A., Coulombe, S., and Sherk, A. (2012). *Methods for constructing life tables for Canada, provinces and territories*. Statistics Canada.

Mathers, C. (1991). *Health expectancies in Australia 1981 and 1988*. Canberra: Australian Institute of Health.

Molla, M.T., Wagener, D.K., and Madans, J.H. (2001). Summary measures of population health: Methods for calculating healthy life expectancy. *Healthy People 2010, Statistical Notes* (21): 1–11.

MSSS (2011). *Pour guider l'action: Portrait de santé et de ses régions, les statistiques*. Québec: Ministère de la Santé et des Services sociaux en collaboration avec l'Institut national de santé publique du Québec et l'Institut de la statistique du Québec.

Page, A., Begg, S., Taylor, R., and Lopez, A.D. (2007). Global comparative assessments of life expectancy: The impact of migration with reference to Australia. *Bulletin of the World Health Organization* 85(6): 474–481.

Pampalon, R., Hamel, D., Gamache, P., Philibert, M.D., Raymond, G., and Simpson, A. (2012). An area-based material and social deprivation index for public health in Québec and Canada. *Canadian Journal of Public Health* 103(Suppl 2): S17–S22.

Pelletier, G. (2005). La population du Québec par territoire des centres locaux de services communautaires, par territoire des réseaux locaux de services et par région sociosanitaire de 1981 à 2026. Québec : Santé et Services sociaux Québec.
http://www.santecom.qc.ca/BibliothequeVirtuelle/MSSS/2550450833.pdf.

Robine, J.M., Romieu, I., and Cambois, E. (1999). Health expectancy indicators. *Bulletin of the World Health Organization* 77(2): 181–185.

Salomon, J.A., Wang, H., Freeman, M.K., Vos, T., Flaxman, A.D., Lopez, A.D., and Murray, C.J. (2012). Healthy life expectancy for 187 countries, 1990–2010: A systematic analysis for the Global Burden Disease Study 2010. *Lancet* 380(9859): 2144–2162. doi:10.1016/S0140-6736(12)61690-0.

Scherbov, S. and Ediev, D. (2012). Significance of life table estimates for small populations: Simulation-based study of standard errors. *Demographic Research* 24(22): 527–550. doi:10.4054/DemRes.2011.24.22.

Schoen, R. (1978). Calculating life tables by estimating Chiang's a from observed rates. *Demography* 15(4): 625–635. doi:10.2307/2061212.

Shyrock, H.S. and Siegel, J.S. (1976). *The methods and materials of demography*. Abridged edition. New York: Academic Press.

Siegel, J.S. (2012). The life table. In: Siegel, J.S. *The demography and epidemiology of human health and aging*. New York: Springer: 135–216. doi:10.1007/978-94-007-1315-4_4.

Silcocks, P.B.S., Jenner, D.A., and Reza, R. (2001). Life expectancy as a summary of mortality in a population: Statistical considerations and suitability for use by health authorities. *Journal of Epidemiology and Community Health* 55: 38–43. doi:10.1136/jech.55.1.38.

Statistics Canada (2006). Census of population. http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3901&lang=en&db=imdb&adm=8&dis=2.

Stratton, J., Mowat, D.L., Wilkins, R., and Tjepkema, M. (2012). Income disparities in life expectancy in the City of Toronto and Region of Peel, Ontario. *Chronic Diseases and Injuries in Canada* 32(4): 208–215.

Sullivan, D.F. (1971). A single index of mortality and morbidity. *HSMHA Health Report* 86(4): 347–354. doi:10.2307/4594169.

Thatcher, A.R., Kannisto, V., and Vaupel, J.W. (1998). The force of mortality at ages 80 to 120. In: Jeune, B. and Vaupel, J.W. (eds.). *Odense Monographs on Population Aging* 5. Odense: Odense University Press.

Ting, L., Yang, C.Y., and Anderson, J.J. (2013). Mortality increase in late-middle and early-old age: Heterogeneity in death processes as a new explanation. *Demography* 50(5): 1563–1591. doi:10.1007/s13524-013-0222-4.

Toson, B. and Baker, A. (2003). *Life expectancy at birth: Methodological options for small populations*. Office for National Statistics.

Vaupel, J.W., Manton, K.G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16(3): 439–454. doi:10.2307/2061224.

Ver Hoef, J.M. (2012). Who invented the delta method? *The American Statistician* 66(2): 124–127. doi:10.1080/00031305.2012.687494.

Von Gaudecker, H.M. and Scholz, R.D. (2007). Differential mortality by lifetime earnings in Germany. *Demographic Research* 17(4): 83–108. doi:10.4054/DemRes.2007.17.4.

Wilmoth, J.R., Andreev, D., Jdanov, D., and Glei, D.A. (2007). *Methods protocol for the human mortality database*. Available at http://www.mortality.org/Public/Docs/MethodsProtocol.pdf.

Wilmoth, J.R. and Lundstrom, H. (1996). Extreme longevity in five countries: Presentation of trends with special attention to issues of data quality. *European Journal of Population* 12(1): 63–93. doi:10.1007/BF01797166.

Appendix 1: Detailed derivation of variance formulae

In the following sections, formulae for LE variance are derived for the variance models examined in this study. Symbols for life table variables and indices follow the convention used in Chiang, which may differ slightly from standard demographic conventions, in order to facilitate comparison between the standard Chiang variance and the variance formulae of this study. All formulae are generally applicable to the variance of life expectancy from age '$\propto$'. Thus for life expectancy from birth as presented in the main text, $\propto= 0$.

1. The standard Chiang variance model

The standard Chiang variance formula is shown below:

$$\sigma_{e_\alpha}^2 = \sigma_{Ch_\propto}^2 = \sum_{x=\alpha}^{w-1} p_{\propto x}^2 [(1 - a_x)n_x + e_{x+1}]^2 \times \sigma_{p_x}^2, \tag{A1.1}$$

where $p_{\propto x}$ is the survival probability from age interval $\propto$ to age interval x, a_x is the fraction of age interval lived for age interval x, n_x is the width (in years) of the age interval, e_{x+1} is the life expectancy from age $x + 1$, and $\sigma_{p_x}^2 = q_x^2(1 - q_x)/D_x$ is the (binomial) variance of the survival probability p_x for the xth age interval.

For comparability with the other variance models in this study, the expression of the standard Chiang variance in terms of the variance of age-specific death counts is readily obtained: $p_x = 1 - D_x/P_x^-$ where P_x^- is the population at the start of age interval x and is considered a constant with no statistical variability; thus $\sigma_{p_x}^2 = \sigma_{D_x}^2/P_x^{-2} = \sigma_{D_x}^2(q_x/D_x)^2$. Direct substitution into Equation A1.1 then yields:

$$\sigma_{e_\alpha}^2 = \sigma_{Ch_\propto}^2 = \sum_{x=\alpha}^{w-1} p_{\propto x}^2 [(1 - a_x)n_x + e_{x+1}]^2 \times \sigma_{D_x}^2 \left(\frac{q_x}{D_x}\right)^2, \tag{A1.1a}$$

Overall, it can be seen that the Chiang LE variance is a weighted sum of the variance contributions of each age interval, with no contribution from the last (w^{th}) interval. The detailed derivation of the standard Chiang LE variance using the delta method approach can be found in Chiang (1984). This derivation and the associated notation form the basis of the derivation of the adjusted and extended variance models shown in the following sections.

2. The 'adjusted' Chiang variance model

The derivation of the adjusted Chiang variance model is based closely on derivation of the standard Chiang LE variance and makes use of the same notation.

For the last age interval ($x = w$), the mortality rate or hazard is assumed to be constant in time, which leads to the standard Chiang closure method[13] for the life-years lived in the last age interval:

$$L_w = \frac{l_w}{M_w}$$

As life expectancy from age $\propto$ equals the sum over age intervals of the life-years lived divided by the starting population at age $\propto$ ($e_\propto = \frac{1}{l_\propto}\sum_{x=\propto}^{w} L_x$), the last age interval thus contributes the following additive term to the life expectancy:

$$\frac{1}{\ell_\alpha} L_w = \frac{1}{\ell_\alpha}\frac{l_w}{M_w} = \frac{p_{\alpha w}}{M_w} = \frac{P_w p_{\propto w}}{D_w} \tag{A1.2}$$

The life expectancy variance, accounting for the contribution of the last age interval death count (D_w), can then be expressed using the delta method in terms of the known variances of $p_\propto, p_{\propto+1}, \dots, p_{w-1}, D_w$ while noting that these variables are statistically independent:

$$\sigma_{e_\alpha}^2 = \sum_{x=\propto}^{w-1} \left(\frac{\partial e_\propto}{\partial p_x}\right)^2 \sigma_{p_x}^2 + \left(\frac{\partial e_\propto}{\partial D_w}\right)^2 \sigma_{D_w}^2 \tag{A1.3}$$

Note that the first part of Equation A1.3 is identical to the standard variance equation (Chiang 1984) since the partial derivatives $\left(\frac{\partial e_\propto}{\partial p_x}\right)$ are unaffected by D_w. Therefore Equation A1.3 can be expressed as the sum of the standard Chiang variance and an additive correction term:

$$\sigma_{e_\alpha}^2 = \sigma_{Ch}^2 + \left(\frac{\partial e_\propto}{\partial D_w}\right)^2 \sigma_{D_w}^2 \tag{A1.4}$$

Using Equation A1.2 the partial derivative $\frac{\partial e_\propto}{\partial D_w}$ is:

---

[13] This closure method is also used by the 'actuarial', Greville, Sirken, Keyfitz and Frauenthal, and Schoen LE estimators.

$$\frac{\partial e_\alpha}{\partial D_w} = \frac{\partial}{\partial D_w}\left(\frac{P_w p_{\alpha w}}{D_w}\right) = -\frac{P_w p_{\alpha w}}{D_w{}^2} \tag{A1.5}$$

Assuming a Poisson distribution for the number of deaths in the last age interval, the variance of the death count in the last age interval is:

$$\sigma^2_{D_w} = D_w \tag{A1.6}$$

Therefore the variance contribution term for the last age interval becomes:

$$\left(\frac{\partial e_\alpha}{\partial D_w}\right)^2 \sigma^2_{D_w} = \left(\frac{P_w p_{\alpha w}}{D_w{}^2}\right)^2 (D_w) = \frac{P_w^2 p_{\alpha w}^2}{D_w^3} = \frac{p_{\alpha w}^2}{M_w^3 P_w} \tag{A1.7}$$

The final expression for the adjusted Chiang variance is thus:

$$\sigma^2_{e_\alpha} = \sigma^2_{LE.adj} = \sigma^2_{Ch_\alpha} + \frac{p_{\alpha w}^2}{M_w^3 P_w} \tag{A1.8}$$

For life expectancy from birth $\alpha = 0$, and the adjusted Chiang variance is written:

$$\sigma^2_{e_0} = \sigma^2_{LE.adj} = \sigma^2_{Ch} + \frac{p_{0w}^2}{M_w^3 P_w} \tag{A1.9}$$

3. The adjusted Chiang variance model with error in population counts

A delta method expansion, accounting for the additional contribution of population error in the last age interval, results in an additional variance contribution term to the adjusted Chiang variance:

$$\sigma^2_{e_\alpha} = \sum_{x=\alpha}^{w-1} \left(\frac{\partial e_\alpha}{\partial p_x}\right)^2 \sigma^2_{p_x} + \left(\frac{\partial e_\alpha}{\partial D_w}\right)^2 \sigma^2_{D_w} + \left(\frac{\partial e_\alpha}{\partial P_w}\right)^2 \sigma^2_{P_w}$$

$$\therefore \sigma^2_{e_\alpha} = \sigma^2_{adj} + \left(\frac{\partial e_\alpha}{\partial P_w}\right)^2 \sigma^2_{P_w} \tag{A1.10}$$

Bourbeau and Lebel (2000) (Table 3) provided estimates of the population error in Canadian data for the census years 1971, 1976, 1981, 1986, and 1991, for the age interval 90+, using the extinct generation method; both over-estimation and under-estimation of population counts were found. Based on these data, a +/– 5% proportional error in the population counts for the age interval 90+ was judged to be a reasonable approximation. Interpretation of this error estimate as a 95% confidence interval for a normally distributed variate leads to the following estimate for the variance of the last age interval population count:

$$2\sigma_{Pw} \approx 0.05 P_w \tag{A1.11}$$

$$\therefore \sigma_{Pw}^2 \approx \left(\frac{0.05}{2}\right)^2 P_w^2 \tag{A1.12}$$

Using equation A1.2, the partial derivative weight for the population count variance is:

$$\left(\frac{\partial e_\infty}{\partial P_w}\right)^2 = \left(\frac{\partial}{\partial P_w}\left(\frac{P_w p_{\infty w}}{D_w}\right)\right)^2 = \left(\frac{p_{0w}}{D_w}\right)^2 \tag{A1.13}$$

Therefore the incremental variance contribution due to population error is:

$$\left(\frac{\partial e_\alpha}{\partial P_w}\right)^2 \sigma_{Pw}^2 = \left(\frac{P_w p_{\alpha w}}{D_w}\right)^2 \left(\frac{0.05}{2}\right)^2 \tag{A1.14}$$

For life expectancy from birth, $\propto= 0$ and the extended variance formula accounting for population error becomes:

$$\sigma_{e_0}^2 = \sigma_{LE.adj.pop}^2 = \sigma_{Ch}^2 + \frac{p_{0w}^2}{M_w^3 P_w} + \left(\frac{P_w p_{0w}}{D_w}\right)^2 \left(\frac{0.05}{2}\right)^2 = \sigma_{adj}^2 + \left(\frac{P_w p_{0w}}{D_w}\right)^2 \left(\frac{0.05}{2}\right)^2 \tag{A1.15}$$

4. The adjusted Chiang variance model in the presence of overdispersion

A natural parameterization (Casella and Berger 2002; DeGroot 1986) of the negative binomial distribution is as follows:

$$D_x \sim NB(P_x^+, q_x), \tag{A1.16}$$

where the age-specific death counts are interpreted as the result of a sequence of Bernoulli trials such that D_x deaths (successes) occur with probability q_x for a given number, P_x^+, of survivors (failures). P_x^+ is thus the number of survivors at the end of age interval x, and q_x is the probability of death for the age interval which can be written $D_x/(P_x^+ + D_x)$.

Under this parameterization, standard formulae (Casella and Berger 2002; DeGroot 1986) for the negative binomial distribution can be used to express the mean and variance of D_x in the following terms:

$$E(D_x) = \frac{q_x P_x^+}{(1 - q_x)} = D_x, \text{ and} \tag{A1.17}$$

$$Var(D_x) = \sigma_{Dx.NB}^2 = \frac{q_x P_x^+}{(1 - q_x)^2} = D_x \left(\frac{P_x^-}{P_x^- - D_x} \right) = D_x \left(\frac{P_x^-}{P_x^+} \right), \tag{A1.18}$$

where $P_x^- = P_x^+ + D_x$ and represents the number of survivors at the start of age interval x.

For age intervals other than the last $(x < w)$, it can be shown that the negative binomial variance $\sigma_{Dx.NB}^2$ is equal to the corresponding non-overdispersed binomial variance, σ_{Dx}^2, multiplied by an inflation factor:

$$\sigma_{Dx}^2 = P_x^- q_x (1 - q_x) = D_x \left(1 - \frac{D_x}{P_x^-} \right) = D_x \left(\frac{P_x^- - D_x}{P_x^-} \right) = D_x \left(\frac{P_x^+}{P_x^-} \right) \tag{A1.19}$$

$$\therefore \sigma_{Dx.NB}^2 = \sigma_{Dx}^2 \left(\frac{P_x^-}{P_x^+} \right)^2 \tag{A1.20}$$

For the last age interval $(x = w)$, $\sigma_{Dx}^2 = D_x$, and therefore

$$\sigma^2_{Dw.NB} = \sigma^2_{Dw} \left(\frac{P_x^-}{P_x^+} \right) \qquad\qquad (A1.21)$$

expresses the degree of variance inflation. Note that for the last age interval, D_x represents number of deaths that occur in this age interval over an 'observation' period of 5 years (= the number of years of data aggregation). Thus P_x^+ represents the number of survivors aged 90+ at the end of this observation period. Insertion of equations A1.20 and A1.21 into the equations for the Chiang or adjusted Chiang variance equations can then be done to readily obtain the overdispersed versions of these two variance models (e.g., Equation 4 of the manuscript).

Finally, it should be noted that the parameters for the negative binomial distribution represent a specific though standard and widely used choice; the actual degree of heterogeneity in real populations is unknown and difficult to estimate. However, the parameterization used results in a variance inflation effect that increases with age and is consistent with the expected increases in frailty in mortality rates due to survivorship effects (Vaupel, Manton, and Stallard 1979), and is thus a reasonable and representative model for heterogeneity.

Appendix 2: Scaling trends in $\Delta\sigma_{adj.Ch}$ vs. a range of life table parameters

Scatterplots of $\Delta\sigma_{adj.Ch}$ versus a range of life table parameters were done to investigate scaling trends and the possibility of thresholds that could be used to classify or identify population strata with elevated last age interval variance contributions. In addition to P_{ttl} (described in the main text), the following seven parameters were examined: M_w (the mortality rate of the last age interval), p_{0w} (the probability of survival from age 0 to w), P_w (the population size in person–years of the last age interval), D_w (the death count of the last age interval), LE (the life expectancy from birth), $\sigma_{LE.Ch}$ (the magnitude of the LE standard error), and the mean population age. M_w, p_{0w}, and P_w correspond to terms present in the adjusted variance correction term (cf. Equation 2), D_w combines both M_w and P_w, LE and $\sigma_{LE.Ch}$ represent basic life table summary statistics, while the mean age represents a proxy for the age structure of the population.

Results are shown in Figures A-2.1a–A-2.1g below. In these figures, crosses, triangles, and gray circles represent results for the provincial, regional, and intra-regional strata respectively, while the .02 and 0.1 thresholds for $\Delta\sigma_{adj.Ch}$ are indicated by dashed lines for reference. $\Delta\sigma_{adj.Ch}$ generally increases with decreasing M_w and P_w and with increasing p_{0w} (Figures A-2.1a–A-2.1c), as would be expected from the equation for the variance contribution term of the last age interval (Equation 2). In particular, scaling trends with P_w are similar to those with P_{ttl} but more variable: for this reason, the latter was chosen for presentation purposes in the main text. Scaling trends with D_w (Figure A-2.1d) are quite similar to those of P_w and P_{ttl}. Increasing trends can additionally be seen in $\Delta\sigma_{adj.Ch}$ vs. LE and $\sigma_{LE.Ch}$ (Figures A-2.1e and A-2.1f). In particular, trends with $\sigma_{LE.Ch}$ are virtually identical to those with P_{ttl} shown in the main text (Figure 1a) due to the close correspondence between the LE standard error and total population size. Finally, Figure A-2.1g shows no clear scaling trend with mean population age. For all cases, however, accurate classification of strata is clearly not possible due to the variability in $\Delta\sigma_{adj.Ch}$ over and above the general scaling trends that exist. Therefore, the mathematical correction of the LE variance, using the readily accessible correction term provided in this manuscript, is recommended to account for the variance contribution of the last age interval in all calculations.

Figure A-2.1: $\Delta\sigma_{adj.Ch}$ plotted against (a) M_w, the mortality rate of the last age interval, (b) p_{0w}, the survival probability from age 0 to w, (c) P_w, the population count of the last age interval, (d) D_w, the death count of the last age interval, (e) LE, the life expectancy from birth, (f) $\sigma_{LE.Ch}$, the Chiang standard error of the life expectancy, and (g) the mean age of the population.
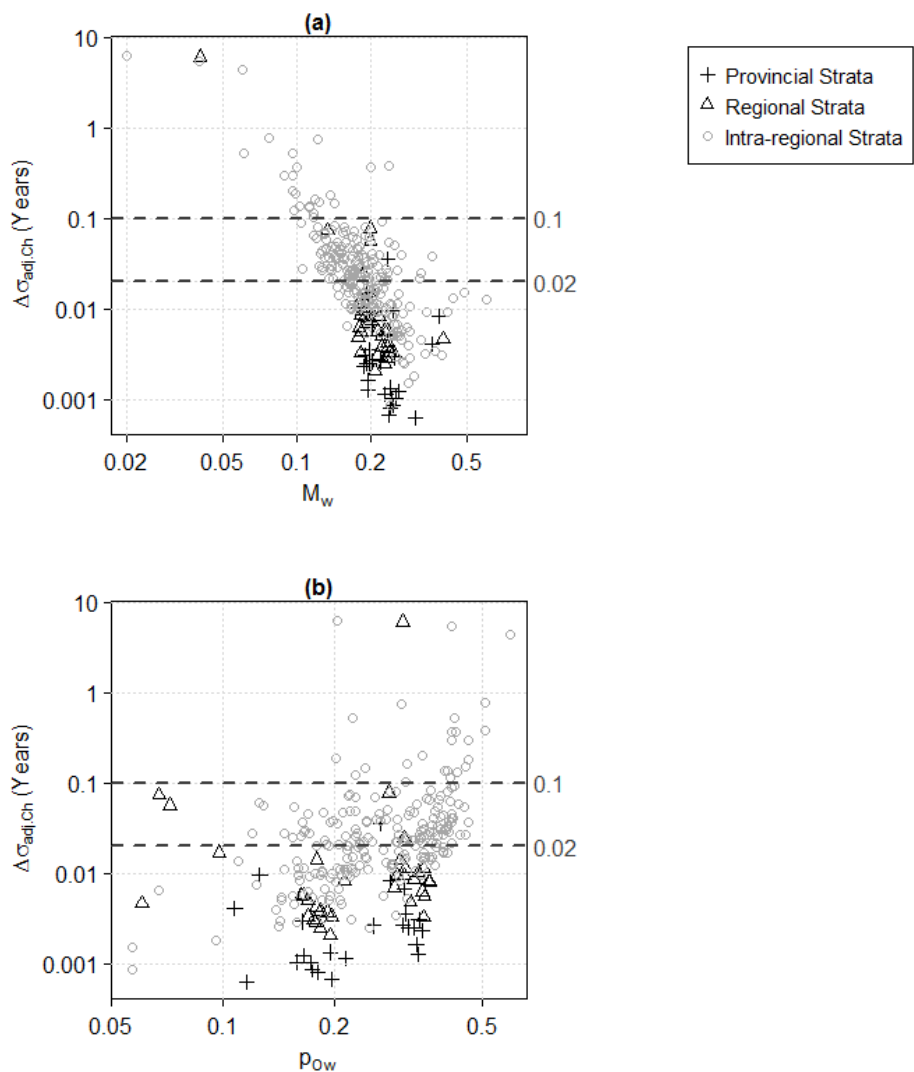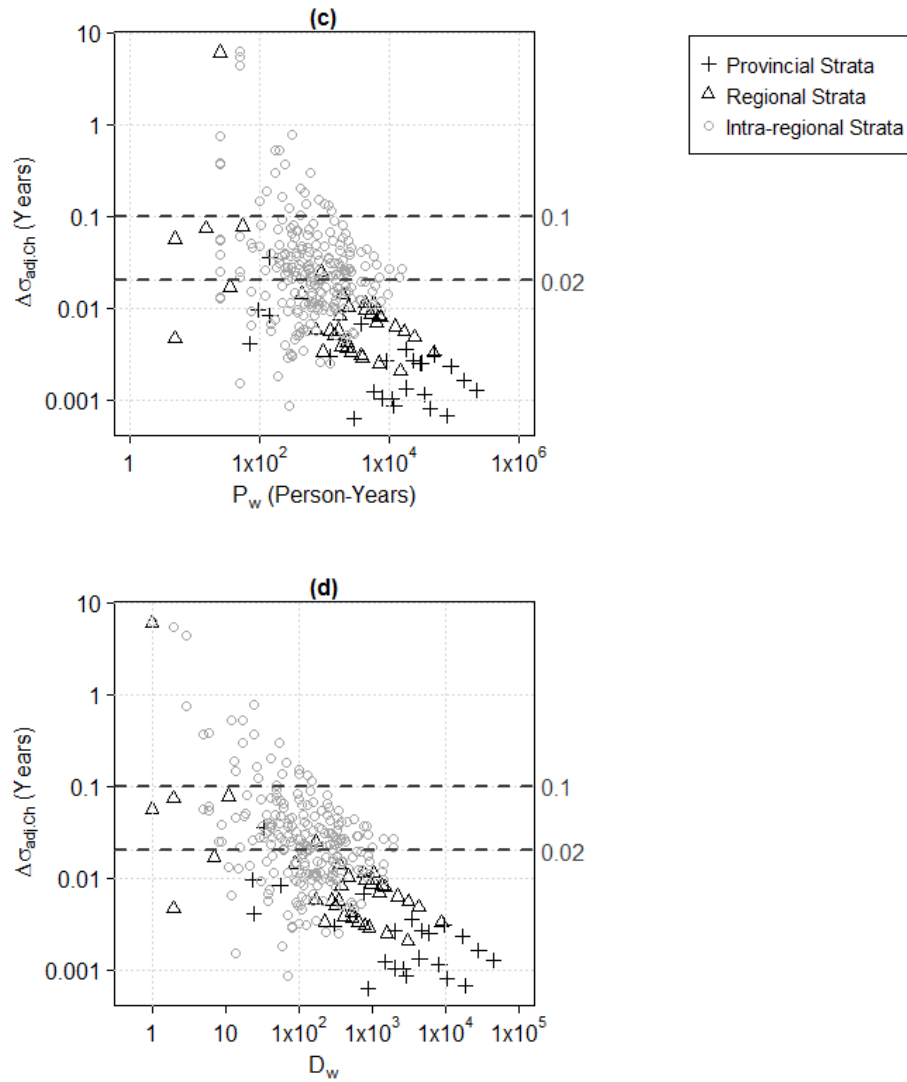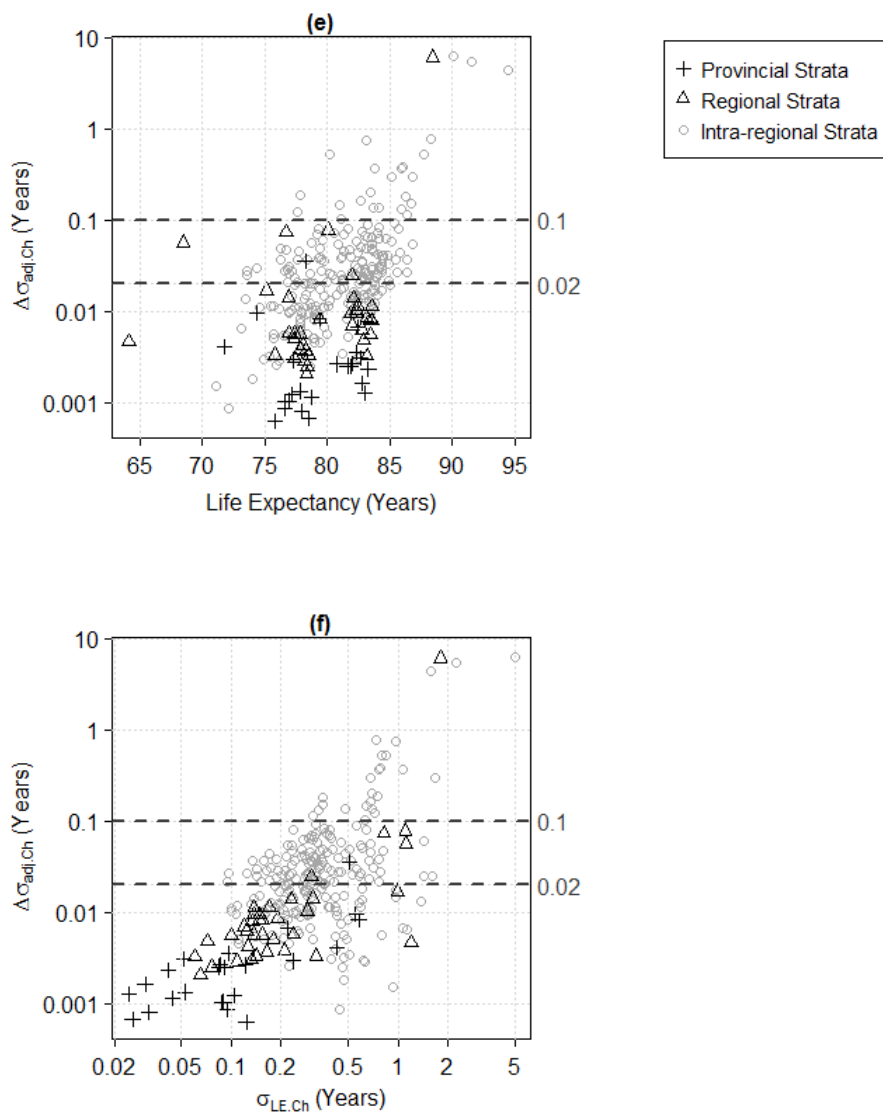
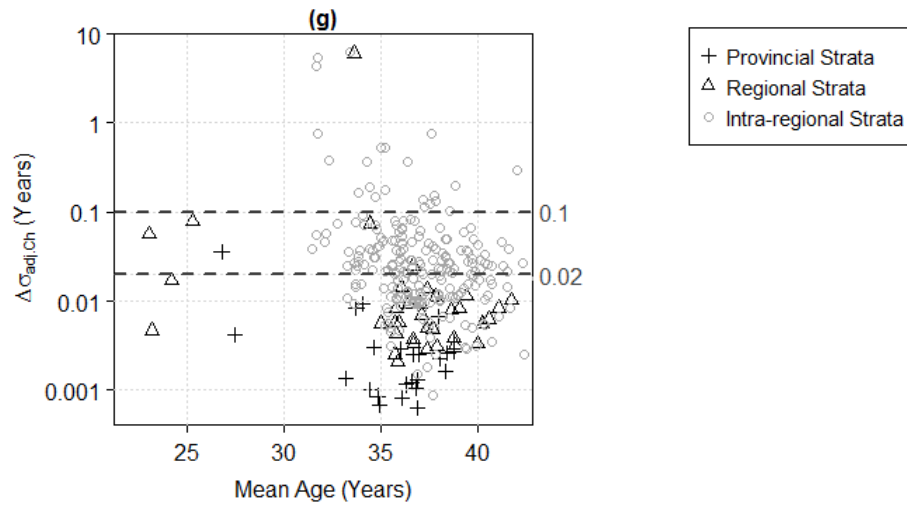**Figure A-2.1:   (Continued)**

Figure A-2.1: **(Continued)**

Figure A-2.1: **(Continued)**

Appendix 3: The contribution of population error to the LE variance – additional analyses and validation studies

1a) Comparison of the 'sensitivity factors' of the last age interval death and population count variances

The delta method expression of the LE variance, accounting for the last age interval death and population count variance, is as follows:

$$\sigma^2_{LE.adj.pop} \approx \sum_{x=0}^{w-1} \left(\frac{\partial LE}{\partial D_x}\right)^2 \sigma^2_{D_x} + \left(\frac{\partial LE}{\partial D_w}\right)^2 \sigma^2_{D_w} + \left(\frac{\partial LE}{\partial P_w}\right)^2 \sigma^2_{P_w}, \tag{A3.1}$$

where the partial derivative multipliers can be interpreted as the 'degree of sensitivity' of the LE variance to each of the component variances.

Noting that the first term in equation A3.1 is simply the standard Chiang variance, and also that the last age interval contributes the term $\frac{L_w}{\ell_0} = \frac{P_w}{D_w} p_{0w}$ to the LE, equation A3.1 becomes:

$$\sigma^2_{LE.adj.pop} \approx \sigma^2_{LE.Ch} + \left[\frac{\partial}{\partial D_w}\left(\frac{P_w}{D_w} p_{0w}\right)\right]^2 \sigma^2_{D_w} + \left[\frac{\partial}{\partial P_w}\left(\frac{P_w}{D_w} p_{0w}\right)\right]^2 \sigma^2_{P_w}$$

$$\approx \sigma^2_{LE.Ch} + \left[\frac{P_w p_{0w}}{D_w^2}\right]^2 \sigma^2_{D_w} + \left[\frac{p_{0w}}{D_w}\right]^2 \sigma^2_{P_w} \tag{A3.2}$$

Thus the ratio of the population count to death count sensitivity factors is found to be equal to the square of the last age interval mortality rate:

$$\left[\frac{p_{0w}}{D_w}\right]^2 / \left[\frac{P_w p_{0w}}{D_w^2}\right]^2 = M_w^2 \tag{A3.3}$$

It can be noted that this is a general result that is independent of any distributional assumptions for P_w and D_w, as well as being independent of any other population characteristic or parameter. While M_w can vary over populations it is invariably $\ll 1$, explaining the reduced impact of population error on the overall LE variance. In the current study, at the median (inter-quartile range) value of M_w of 0.19 (0.16–0.23), the sensitivity factor of the population count is less than 4% (3%–5%) that of the death count.

1b) Comparison of the magnitudes of the last age interval death and population count variance contributions for the specific case of a 5% proportional error

Substitution of estimated expressions for the variances themselves permits comparison of the magnitudes of the last age interval death and population count variance contributions. Using a Poisson distribution for D_w (cf. Appendix 1, Section 1) and a 5% proportional error for P_w (cf. Appendix 1, Section 3) yields:

$$\sigma^2_{LE.adj.pop} \approx \sigma^2_{LE.Ch} + \left[\frac{P_w p_{0w}}{D_w^2}\right]^2 D_w + \left[\frac{p_{0w}}{D_w}\right]^2 \left(\frac{0.05}{2}\right)^2 P_w^2$$

$$\approx \sigma^2_{LE.Ch} + \frac{P_w^2 p_{0w}^2}{D_w^3} + \frac{P_w^2 p_{0w}^2}{D_w^4}\left(\frac{0.05}{2}\right)^2 \tag{A3.4}$$

Therefore the ratio of the population count to death count variance contributions can be expressed in terms of the last age interval mortality death count:

$$\left(\frac{P_w p_{\alpha w}}{D_w}\right)^2 \left(\frac{0.05}{2}\right)^2 \bigg/ \frac{P_w^2 p_{0w}^2}{D_w^3} = 0.00125 \times D_w \tag{A3.5}$$

Thus the variance contribution of the population count is comparable to that of the death count when $0.00125 \times D_w \gtrsim 1$ or $D_w \gtrsim 1600$. This condition is satisfied only in very large populations where overall variance contribution is small. In the current dataset $D_w \gtrsim 1600$ corresponds to $P_{ttl} \gtrsim 1.2 \times 10^6$.

2. Sensitivity of the LE variance to assumptions of confidence interval coverage in the last age interval population count error

The 5% proportional error suggested by Bourbeau and Lebel (Bourbeau and Lebel 2000) could well correspond to a confidence interval coverage different from the 95% coverage that was assumed. It can be demonstrated, however, that the variance contributed by error in the last age interval population count is relatively insensitive to different levels of assumed confidence interval coverage.

Equation 3 for the LE variance including population error can be written in terms of a general coverage parameter or critical value z:

$$\sigma^2_{LE.adj.pop} \approx \sigma^2_{LE.Ch} + \left(\frac{p_{0w}}{M_w}\right)^2 \left(\frac{0.05}{z}\right)^2,$$

(A3.6)

where z≈2 was used for the 95% confidence interval coverage analyzed in the manuscript main text. Based on a normal distribution model (reasonable, as the empirical data show that the error oscillates or appears distributed symmetrically around a central value), values of z≈1.6 and 2.6 can be taken to correspond approximately to 90% and 99% coverage values.

Figure A-3.1a-c shows $\Delta\sigma_{pop.adj}$ calculated for critical z-values of 1.6, 2, and 2.6 over the datasets described in the manuscript. It can be seen that although $\Delta\sigma_{pop.adj}$ increases slightly with decreasing z (or coverage) it remains small overall, with values never exceeding 0.04 over all datasets. Expressed alternatively, there are no strata for which $\Delta\sigma_{pop.adj}$>0.02 when z=2.6: this rises to 5 strata when z=2 and to just 12 when z=1.6. In conclusion, the variance contribution of population error in the last age interval to the total LE variance is relatively insensitive to different coverage values. Mathematically, this is explained by the insensitivity of the population variance contribution term to different values of z near 2 and the fact that the overall magnitude of the contribution of population error is relatively small (due to its small sensitivity factor, as was shown above). Thus the overall finding of our study that the contribution of last age interval population error can be neglected is robust with respect to assumptions of coverage.

Figure A-3.1: $\Delta\sigma_{pop.adj}$ **for coverage parameters (a) z = 1.6, (b) z = 2, and (c) z = 2.6**
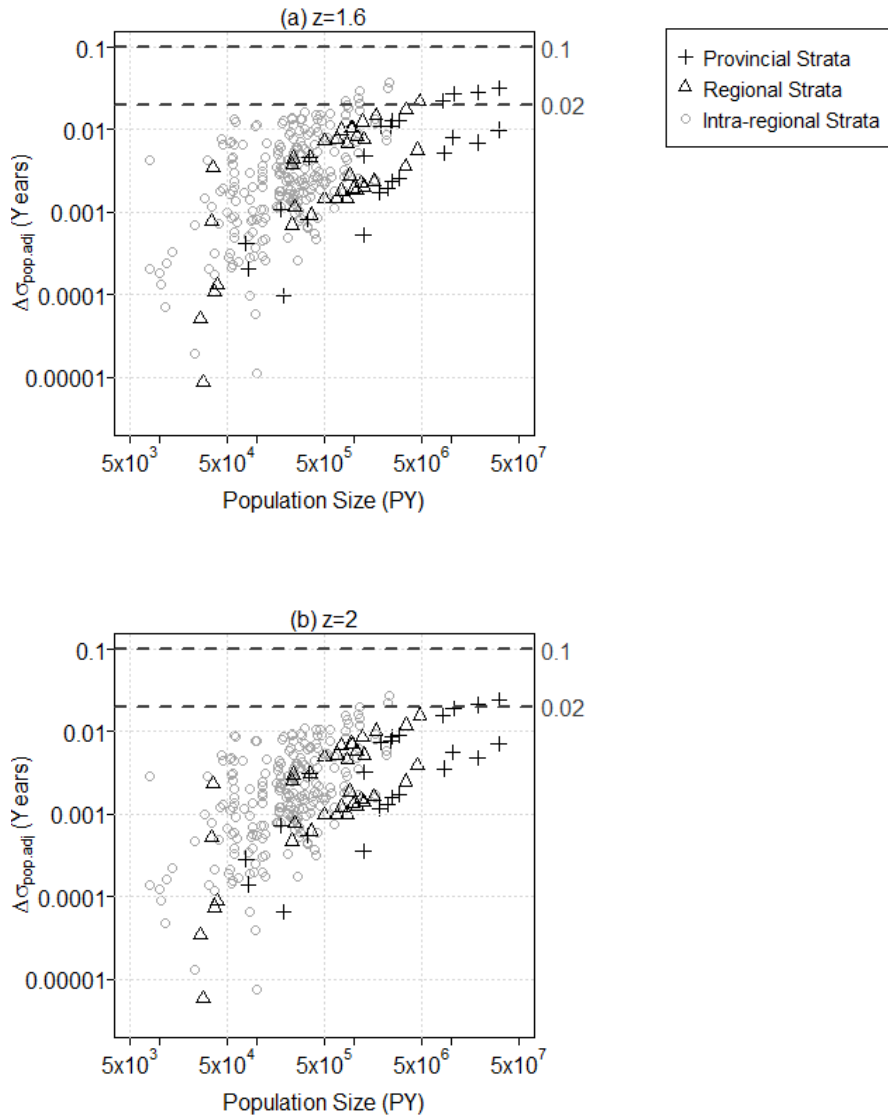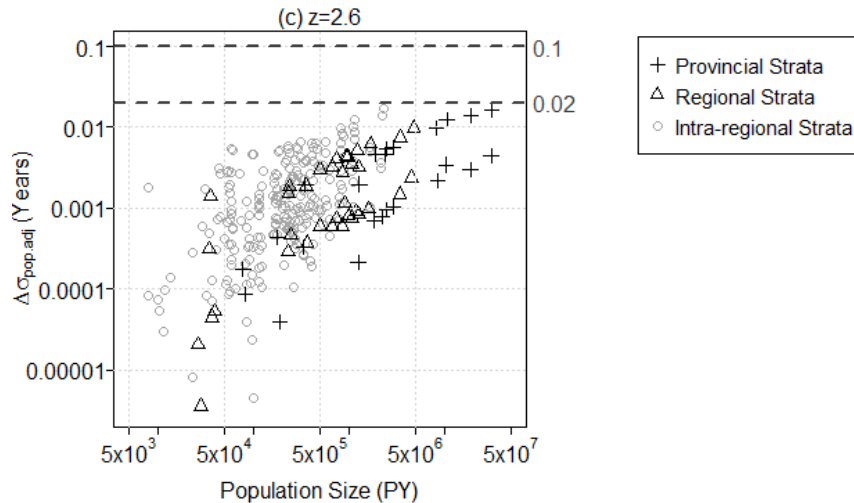
Figure A-3.1:   (Continued)

3. The effect of a corresponding 'opposing' error occurring in younger age intervals

Age misreporting is likely a primary mechanism behind the last age interval population count error. This type of error in principle should entail a corresponding opposite error at younger age groups, as individuals are misclassified by age but not excluded from population counts. However, the impact of 'opposing population error' that might occur at younger age groups should be much smaller than that of the open age group due to the larger total population sizes at these younger age groups. For example, over all strata considered in the current study, the median (inter-quartile range) ratio of the population size of the second-to-last age group to the last open age group is $P_{85-89}/P_{90} \approx 2.2$ (1.9 − 2.8), while that of the third-to-last age group is $P_{80-85}/P_{90} \approx 4.8$ (3.7 − 6.8). Thus the impact of a 5% error in the population in the last age group, distributed over these last two age groups is expected to be much smaller.

Any corresponding population error at younger age groups will furthermore act in an opposite direction and thus reduce the overall variance contribution of population error. As a result, the variance contribution estimate based on the last age interval population error can be considered as a conservative upper limit to the actual total population error. Consequently, the finding of our study that the contribution of population error is small and can be neglected remains robust and valid.

The above two effects can be demonstrated mathematically by considering the specific and not unreasonable case where error in last age interval (90+ or wth) population counts results in a corresponding opposite error in the second-to-last (85–89 or (w−1)th) age interval. The delta method expression for the LE variance including the population count error of the (w-1)th age interval can be written as follows:

$$\sigma^2_{LE.adj} \approx \sum_{x=0}^{w} \left(\frac{\partial LE}{\partial D_x}\right)^2 \sigma^2_{D_x} + \left(\frac{\partial LE}{\partial P_w}\right)^2 \sigma^2_{P_w} + \left(\frac{\partial LE}{\partial P_{w-1}}\right)^2 \sigma^2_{P_{w-1}} \\ + 2\left(\frac{\partial LE}{\partial P_w}\right)\left(\frac{\partial LE}{\partial P_{w-1}}\right)\sigma_{P_w}\sigma_{P_{w-1}}\rho_{P_w,P_{w-1}}$$

(A3.7)

where the last term represents the effect of correlation between the population counts of the wth and (w-1)th age intervals and $\rho_{P_w,P_{w-1}}$ represents the correlation coefficient between P_w and P_{w-1}. Since for each possible deviation in P_w there is an equal and opposite deviation in P_{w-1}, $\rho_{P_w,P_{w-1}} = -1$ and $\sigma_{P_{w-1}} = \sigma_{P_w}$. Therefore,

$$\sigma^2_{LE.adj.pop} \approx \sum_{x=0}^{w} \left(\frac{\partial LE}{\partial D_x}\right)^2 \sigma^2_{D_x} + \left(\frac{\partial LE}{\partial P_w}\right)^2 \sigma^2_{P_w} + \left(\frac{\partial LE}{\partial P_{w-1}}\right)^2 \sigma^2_{P_w} - 2\left(\frac{\partial LE}{\partial P_w}\right)\left(\frac{\partial LE}{\partial P_{w-1}}\right)\sigma^2_{P_w}$$

$$= \sum_{x=0}^{w} \left(\frac{\partial LE}{\partial D_x}\right)^2 \sigma^2_{D_x} + \left(\frac{\partial LE}{\partial P_w} - \frac{\partial LE}{\partial P_{w-1}}\right)^2 \sigma^2_{P_w}$$

(A3.8)

Comparison with Equation A3.1, and the fact that $\frac{\partial LE}{\partial P_w}$ and $\frac{\partial LE}{\partial P_{w-1}}$ have the same sign, indicates that the effect of population error at a younger age group is to reduce the overall variance contribution due to population error.

As a validation analysis, Equation A3.8 was evaluated over all strata used in the current study. Comparison of these results (Figure A-3.2 below) with those of the main text (Figure 3) confirms the overall decrease in the change in standard error ($\Delta\sigma_{pop.adj}$) that results from an opposing population error in the next-to-last age group.

Figure A-3.2: $\Delta\sigma_{pop.adj}$ **including the contribution of the (w-1)th age interval vs. total population size**
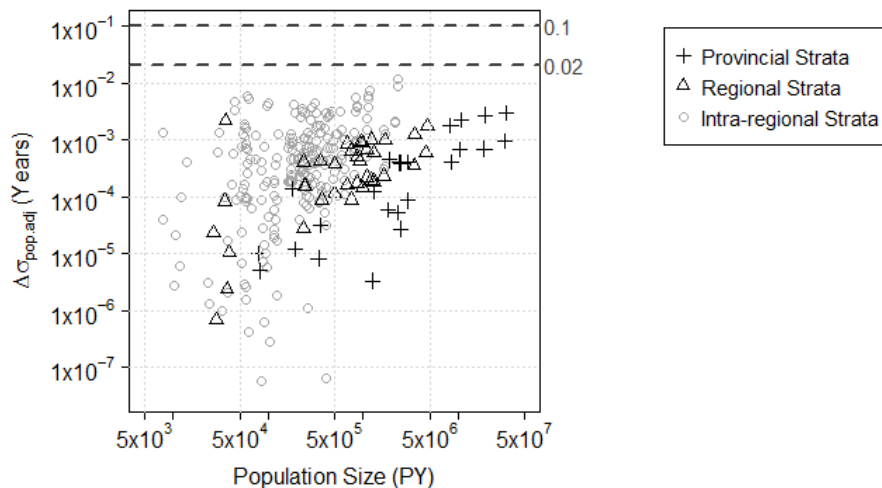
In conclusion, mathematical and empirical validation analyses show that the overall finding of the manuscript, that the contribution of the last age interval population error is comparatively small and can be neglected, is robust and valid.

Appendix 4: Assessment of the impact of the last age interval on the variance of health expectancy (HE) using health-adjusted life expectancy (HALE) as an example

Theory

Health expectancy (HE) is most often estimated using the Sullivan method (Sullivan 1971), in which the number of life-years contributed by each age interval of the life table is weighted by a measure of health status (h_x). Therefore sources of variability in HE generally comprise both the age-specific death counts (D_x) and the age-specific health status weights (h_x). The variance model for health expectancy that correctly includes the effect of the last age interval is thus:

$$\sigma_{HE}^2 = f(D_x, D_w, h_x, h_w | P_x, P_w), \ \ \text{for } x = 0 \dots \text{w-1} \tag{A4.1}$$

Mathers (1991) was the first to extend the delta method approach used by Chiang to derive the currently well-known formula for HE variance. Although this formula included the variance contribution of the health status of the last age interval, the contribution of the last age interval death counts was omitted, consistent with Chiang's original derivation of the LE variance.

Applying a delta method expansion to HE while including the variance contribution of the last age interval death counts yields the following expression for the HE variance:

$$\sigma_{HE.adj}^2 \approx \sum_{x=0}^{w-1} \left(\frac{\partial HE}{\partial D_x}\right)^2 \sigma_{D_x}^2 + \sum_{x=0}^{w} \left(\frac{\partial HE}{\partial h_x}\right)^2 \sigma_{h_x}^2 + \left(\frac{\partial HE}{\partial D_w}\right)^2 \sigma_{D_w}^2$$

$$= \sum_{x=\alpha}^{w-1} p_{\alpha x}^2 [(1-a_x)n_x h_x + e_{x+1}]^2 \times \sigma_{px}^2 + \sum_{x=\alpha}^{w} \left(\frac{L_x}{\ell_x}\right)^2 \times \sigma_{hx}^2 + \frac{\ell_w^2 h_w^2}{M_w^3 P_w}, \tag{A4.2}$$

where the variance contribution of the last age interval death counts, D_w, has been written as a separate term.

The above corrected variance formula can be written as the sum of the standard 'Chiang-Mathers' variance and an additional term due to the omitted last age interval contribution:

$$\sigma^2_{HE} \approx \sigma^2_{HE.Ch.Mat} + \frac{p^2_{\propto w} h^2_w}{M^3_w P_w} \tag{A4.3}$$

It can be seen that the 'standard' Chiang-Mathers variance for HE underestimates the actual variance, but can be corrected with a simple additive term which represents the impact of the last age interval variance contribution.

Methods

Health-adjusted life expectancy (HALE) calculated using the Sullivan method (Sullivan 1971) will be used as a representative example to demonstrate the potential impact of the last age interval contribution to the HE variance. Analyses are done for the regional-level strata only. Health status (h_x) for HALE is represented by the HUI (Health Utilities Index Mark 3) (Horsman et al. 2003), and is a continuous measure[14] with an upper limit of 100% and that can be interpreted as the proportion of life-years spent in good health (Manuel, Schultz, and Kopec 2002; Salomon et al. 2012). HUI is estimated by sex, age group, and region, from national-level survey data. Due to lack of available data, HUI for regions 17 and 18 were imputed using values from region 10, which has a similar population size and is geographically proximate. The difference in corrected vs. uncorrected HE standard errors, $\Delta\sigma_{adj.Ch-Mat} = \sigma_{HE.adj} - \sigma_{HE.Ch-Mat}$, is used to assess the impact of the last age interval variance contribution.

Results

In Figure A-4.1, $\Delta\sigma_{adj.Ch-Mat}$ is shown plotted vs. P_{ttl} over the 36 regional-level strata; dashed lines indicate the 0.02 and 0.1 threshold values. The prevalence of $\Delta\sigma_{adj.Ch-Mat}$ exceeding these two thresholds is presented in Table 4.1. Overall, it can be seen that the last age interval can contribute substantially to the HALE variance. Approximately 11% (4 of 36) of strata exhibit SE.diff higher than 0.02, while 3% (1 of 36) of strata have SE.diff higher than 0.1 years, despite all population strata being well in excess of 5,000 PY in size. In general, elevated variance underestimation in HALE occurs for the same strata as that for LE, though the magnitude of the underestimation is

---

[14] Note that for the HALE metric in particular, h_x is a continuous measure, while for other HE metrics h_x is generally the population prevalence of a dichotomous measure of health status. Equations A3.1–A3.3 are valid for either case.

proportionately smaller due to the additional variance contribution of the health status weight variable (HUI).

Table A-4.2 shows the impact of the standard Chiang variance on statistical tests of HALE. The induced, additional false positive rate for tests of HALE was found to be 2.8% for nominal false positive rates (or significance levels) of 5% and 1% respectively.

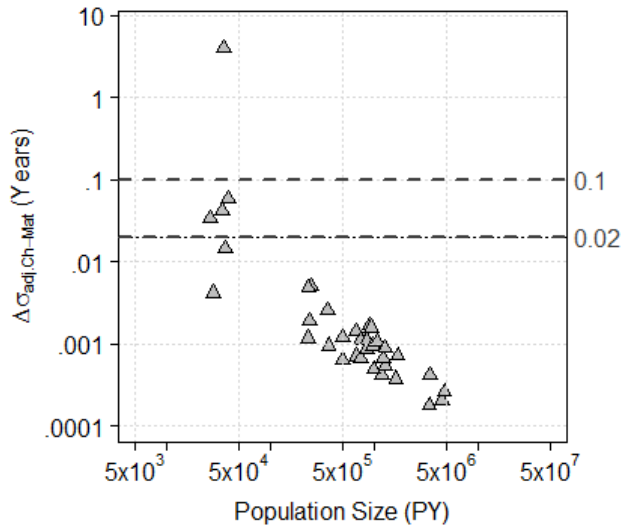**Figure A-4.1:** $\Delta\sigma_{adj.Ch-Mat}$ **vs. population size over regional-level strata**

Table A-4.1: Proportion of strata with elevated $\Delta\sigma_{adj.Ch-Mat}$

Dataset	# of Strata	% $\Delta\sigma_{Ch-Mat.adj} > 0.02$	% $\Delta\sigma_{Ch-Mat.adj} > 0.1$
Regional	36	11.1% (4/36)	2.8% (1/36)

Table A-4.2: Induced false positive rates for $\sigma^2_{HE.Ch-Mat}$ with respect to $\sigma^2_{HE.adj}$

Dataset	# of Statistical Comparisons	False positive rate, α=0.05	False positive rate, α=0.01
Regional	36	2.8% (1/36)	2.8% (1/36)

Conclusions

The application of the adjusted Chiang method to HE has not been previously noted in the scientific literature, and this is an especially pertinent issue given the increasing usage of and preference for HE metrics to quantify and compare population health (Loukine et al. 2012; Molla, Wagener, and Madans 2001; Robine, Romieu, and Cambois 1999; Salomon et al. 2012), and the frequent presentation of LE and HE measures together (Bajekal 2005; Geronimus et al. 2001; Salomon et al. 2012). In the present study, a simple additive variance correction term is derived for HE. Using HALE as an example, it is demonstrated that substantial variance underestimation can occur unless the adjusted Chiang method is used. In particular, cases may readily arise when the variance contribution of the last age interval exceeds that of the health status component and thus dominates the total variance.

Appendix 5: The effect of a reduction in closure age on the last age interval variance contribution

The effect of varying the closure age on the variance contribution of the last age interval was examined by applying a reduced closure age of 85+ years followed by an increased closure age of 95+ years. For each case, the difference in LE standard error with and without the variance contribution of the last age interval, denoted by $\Delta\sigma_{adj.Ch.85}$ and $\Delta\sigma_{adj.Ch.95}$ respectively, was estimated and compared to the baseline closure age result using 90+ years. In the text that follows, $\Delta\sigma_{adj.Ch.90}$ represents that estimated using a closure age of 90 years (which was denoted simply as $\Delta\sigma_{adj.Ch}$ in the main text).

Figure A-5.1 shows $\Delta\sigma_{adj.Ch.85}$ plotted vs. $\Delta\sigma_{adj.Ch.90}$: crosses, triangles, and gray circles represent results for the provincial, regional, and intra-regional strata respectively, while the 1–1 line of equality is indicated by the black dotted line. For virtually all populations, reduction of the life table closure age from 90 to 85 years in fact further increases the variance underestimation by the standard Chiang method, as indicated by the rightward shift in data points with respect to the 1–1 line. This increase in variance contribution can be explained by the effect of the increased probability of survival from age 0 to w (p_{0w}) and reduced mortality rate (M_w), for a last age interval of 85+ vs. 90+ years on the variance correction term (cf. Equation 2).

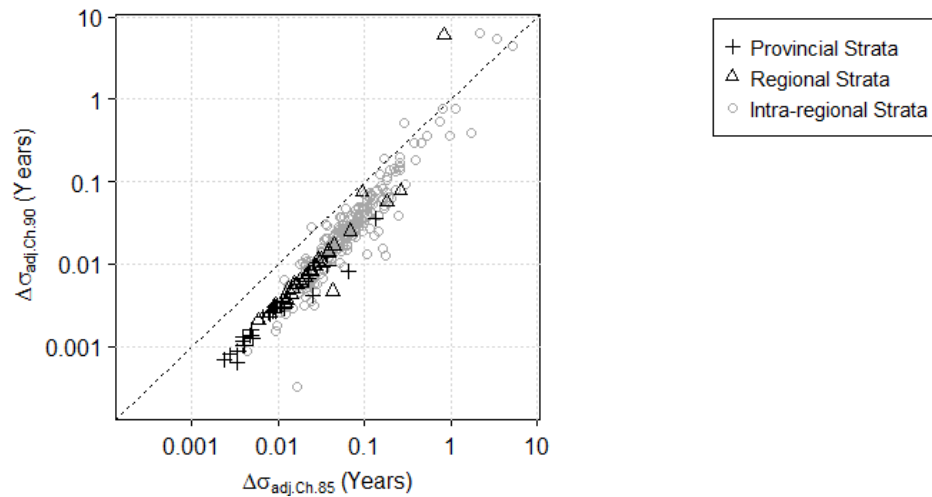**Figure A-5.1:** $\Delta\sigma_{adj.Ch.90}$ vs. $\Delta\sigma_{adj.Ch.85}$

Figure A-5.2: $\Delta\sigma_{adj.Ch.90}$ vs. $\Delta\sigma_{adj.Ch.95}$[15]
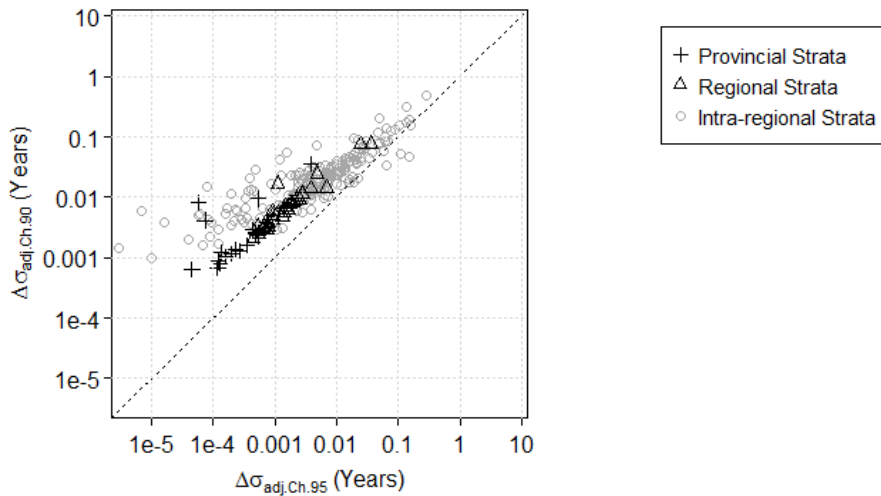

Figure A-5.2 shows $\Delta\sigma_{adj.Ch.90}$ plotted vs. $\Delta\sigma_{adj.Ch.95}$: crosses, triangles, and gray circles represent results for the provincial, regional, and intra-regional strata respectively, while the 1−1 line of equality is indicated by the black dotted line. In this case, increase of the closure age is shown to decrease the variance underestimation by the standard Chiang method, as indicated by the leftward shift in data points with respect to the 1−1 line. The decrease in variance contribution is explained by the decreased probability of survival from age 0 to w (p_{0w}) and increased mortality rate (M_w), for a last age interval of 95+ vs. 90+ years on the variance correction term (cf. Equation 2). Note, however, that the number of excluded strata rises substantially when the closure age is increased to 95+ years, to 18% of sub-regional strata and 8% of regional strata.

In conclusion, reduction of the life table closure age further amplifies the variance contribution of the last age interval and hence the variance underestimation by the standard Chiang method. Increase of the life table closure age diminishes the magnitude of variance underestimation, but leads to a substantial increase in excluded strata, and is not a practical solution for LE estimation at subnational scales.

---

[15] Note that results of Figure A-5.2 were calculated in a later validation analyses,: thus there may be minor shifts in the extracted death and population data used.