Continuous Summary Measures for Changes in the Distribution of Ages at Death

Nadine Ouellette and Robert Bourbeau
Department of Demography, Université de Montréal
Correspondence: nadine.ouellette@umontreal.ca

PLEASE DO NOT CITE WITHOUT THE PERMISSION OF THE AUTHORS
Paper submitted for the Population Association of America (PAA) 2010 Annual Meeting

Background

Since the beginning of the twentieth century, remarkable modifications have occurred in terms of distribution of ages at death in human populations. The increase in the central age at which individuals die and the reduction in the variability of ages at death have been largely documented for several low mortality countries (see Cheung et al. (2005) for a comprehensive list of publications). Up until the 1950s, important mortality gains among infants, children and even young adults occurred in these countries and the compression of their overall age-at-death distribution was consequently strong during that period. Nowadays, this overall compression has slowed down substantially (Wilmoth and Horiuchi, 1999; Yashin et al., 2001), even though important gains have been consistently recorded among older adults for several decades (Jeune and Vaupel, 1995; Kannisto et al., 1994). These findings led researchers to distinguish old-age mortality compression from overall mortality compression and to study changes in the age-at-death distribution at older ages and over the entire age range (or a very broad range) separately.

Inspired by Lexis's concept of normal life durations (Lexis, 1877, 1878), Kannisto (2000, 2001) indeed focused specifically on changes in the distribution of ages at death at older ages and recommended using the adult modal age at death M as a central longevity indicator.

Since this indicator is solely influenced by adult mortality, it is more sensitive to changes occurring among the elderly population than life expectancy at birth for example. To measure variability in old-age mortality, Kannisto notably suggested the standard deviation of individual life durations above M (the root mean square of individual life durations from M, for those individuals that reach M) $SD(M+)$ and the life expectancy at the mode. The results showed that during the second half of the twentieth century, increases in the modal age at death were paralleled by decreasing variability in old-age mortality. In other words, old-age mortality compression was well at work.

Kannisto's approach with emphasis on M was elaborated further in several recent studies (Canudas-Romo, 2008; Cheung et al., 2005, 2008, 2009; Cheung and Robine, 2007; Ouellette and Bourbeau, 2009; Thatcher et al., 2010). Those focusing specifically on changes in the distribution of ages at death at older ages underline that some of today's lowest mortality countries might have taken on new paths where, notably, deaths above M no longer tend to concentrate into a narrower age interval over time (Cheung et al., 2008, 2009; Cheung and Robine, 2007; Ouellette and Bourbeau, 2009). The shifting mortality regime introduced by Kannisto (1996) and further described by Bongaarts (2005) under which adult mortality is assumed to shift to higher ages over time has been identified as a plausible successor. Nevertheless, deviations from this regime have been noticed and call for further research on the topic.

Accordingly, the objective of this paper is to provide a flexible two-dimensional nonparametric approach which leads to smoothed age-at-death distributions and refines our monitoring of changes that have occurred at older ages since 1950. This approach generalizes the one developed by Ouellette and Bourbeau (2009) where calendar years were treated independently from one another. Given the great diversity of historical mortality pathways experienced by today's low mortality countries, we have included the countries of Canada, Denmark, England and Wales, France, Italy, Japan, the Netherlands, Sweden, Switzerland and the United States in our analysis.

Data and methods

For each of the countries selected, data on deaths and exposure to risk starting in 1950 are taken from the Human Mortality Database (HMD 2010). Specifically, for a given country and sex, observed death counts by single year of age i and single calendar year j are arranged into matrix $\boldsymbol{Y} = (y_{ij})$. Similarly, $\boldsymbol{E} = (e_{ij})$ denotes the matrix of exposures to risk and we define $\boldsymbol{M} = \boldsymbol{Y}/\boldsymbol{E}$ as the matrix of observed death rates.

One-dimensional Poisson P-spline smoothing

In previous work (Ouellette and Bourbeau, 2009), we introduced the use of a one-dimensional nonparametric approach based on regression splines to study changes in the distribution of ages at death. For a given year k (subscript k is omitted in this subsection to simplify the notation), we assumed that death counts y_i were a realization of a Poisson distribution with mean $e_i \cdot \mu_i$ and thus obtained the following equality from the Poisson regression model:

$$\log(\boldsymbol{e} \cdot \boldsymbol{\mu}) = \log(\boldsymbol{e}) + \log(\boldsymbol{\mu}).$$

In order to find a smooth estimate for $\boldsymbol{\mu}$, we used the method of P-splines (Eilers and Marx, 1996), which combines the concepts of B-splines and penalized likelihood. The B-splines provide flexibility while the penalty acts on adjacent coefficients of the B-splines to ensure smoothness. Thus, we obtained

$$\log(\hat{\boldsymbol{\mu}}(x)) = \boldsymbol{B}(x)\hat{\boldsymbol{a}},$$

where $\boldsymbol{B}(x)$ is the B-spline basis regression matrix and $\hat{\boldsymbol{a}}$ is the vector of estimated coefficients associated to each B-spline included in the basis. Thus, $\boldsymbol{B}(x)$ has as many columns as there are B-splines in the basis and the values of these B-splines at each age are indexed by row. Note that since infant and child mortality present unique features that would require the use of a smoothing method suited for ill-posed data, mortality and exposure data before age ten were excluded from model estimation.

The corresponding smoothed survival function expressed as

$$\hat{\boldsymbol{S}}(x) = \exp\left(-\int_0^x \hat{\mu}(t)dt\right)$$

can then be calculated using standard numerical techniques. Finally, the smoothed probability density function describing the age-at-death distribution in the given year k corresponds to

$$\hat{\boldsymbol{f}}(x) = \hat{\boldsymbol{\mu}}(x)\hat{\boldsymbol{S}}(x).$$

As opposed to common parametric approaches to model the age pattern of mortality (ex. Gompertz, logistic, Siler, and so on), the method of P-splines does not rely on any rigid theoretical assumptions or modelling structure. Such flexibility leads to a finer expression of the underlying mortality trend over age, as described by the actual data, for every calendar year under study. It consequently improved our ability to monitor alterations in the ages at death distribution over time.

Since mortality change over age and over time is expected to be regular over both ages and years, this paper provides a generalization of this one-dimensional Poisson P-spline smoothing method. Indeed, we introduce the use of a two-dimensional nonparametric smoothing approach (Camarda, 2008, 2009; Currie et al., 2004, 2006; Eilers and Marx, 2002; Eilers et al., 2006) to study recent modifications in the distribution of ages at death in the various low mortality countries selected. Thus, successive calendar years are no longer be treated independently because the additional dimension in this model enables us to take into account the continous nature of mortality progression along ages and over the years simultaneously.

Unlike the well-known age-period-cohort model (Clayton and Schifflers, 1987) and Lee-Carter approaches (Lee and Carter, 1992; Brouhns et al., 2002), ours does not make any rigid assumption about the functional form of the mortality surface. Given the objective of this paper, mortality surfaces are already so informative that imposing a strong model structure on it seems unnecessary.

Two-dimensional Poisson P-spline smoothing

To move from one-dimensional to two-dimentional Poisson P-spline smoothing, one must find a new B-spline basis adapted for two-dimensional regression. Let $\boldsymbol{B}_{age} = \boldsymbol{B}(\boldsymbol{x}_{age})$ denote a regression matrix of B-splines for the ages, and similarly, let $\boldsymbol{B}_{year} = \boldsymbol{B}(\boldsymbol{x}_{year})$ be a regression matrix of B-splines for the years. Then, the new regression matrix $\boldsymbol{B}$ to be used for two-dimesional Poisson P-spline smoothing consists in the following Kronecker product of $\boldsymbol{B}_{age}$ and $\boldsymbol{B}_{year}$:

$$\boldsymbol{B} = \boldsymbol{B}_{year} \otimes \boldsymbol{B}_{age}. \tag{1}$$

As shown in figure 1, the Kronecker product of two one-dimensional B-splines (one along the age dimension and the other along the year dimension) gives rise to an hill. Thus, in equation 1, the Kronecker product of two B-splines basis $\boldsymbol{B}_{age}$ and $\boldsymbol{B}_{year}$ will populate the age-year grid by several overlapping and regularly spaced hills such as the one of figure 1. Indeed, the complete illustration of $\boldsymbol{B}$ (not shown here) includes about 300 overlapping hills providing great flexibility in the modelling process.

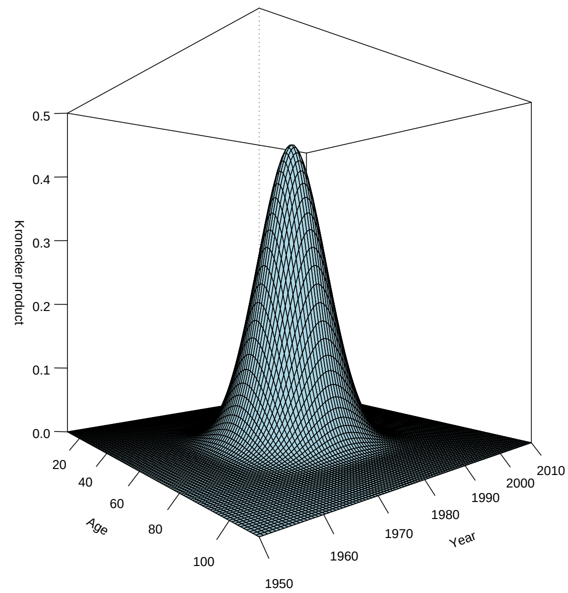


Figure 1: Two-dimensional Kronecker product of two cubic B-splines

Arranging observed deaths and exposures to risk into column vectors $\boldsymbol{y} = \text{vec}(\boldsymbol{Y})$ and $\boldsymbol{e} = \text{vec}(\boldsymbol{E})$ leads to the following expression for the smoothed force of mortality:

$$\log(\hat{\boldsymbol{\mu}}) = \log(\hat{\mathbb{E}}(\boldsymbol{y})) - \log(\boldsymbol{e}) = \boldsymbol{B}\hat{\boldsymbol{a}}.$$

More explicitly, we have

$$\hat{\boldsymbol{\mu}}(\boldsymbol{x}_{age}, \boldsymbol{x}_{year}) = \exp\left((\boldsymbol{B}_{year} \otimes \boldsymbol{B}_{age})\hat{\boldsymbol{a}}\right).$$

For example, figure 2 shows observed and smoothed death rates among Danish males from 1950 to 2007. Another informative way to display mortality rates by age and year is on a mortality surface. Indeed, mortality surfaces provide interpretation at a glance and greatly ease comparisons between countries. Therefore, smoothed mortality surfaces by sex for all countries under study are presented in the Appendix.

Using standard numerical integration techniques, we can compute the smoothed survival function

$$\hat{\boldsymbol{S}}(\boldsymbol{x}_{age}, \boldsymbol{x}_{year}) = \exp\left(-\int_0^{x_{age}} \hat{\boldsymbol{\mu}}(t, \boldsymbol{x}_{year})dt\right).$$

The smoothed density function describing the age-at-death distributions then corresponds to

$$\hat{\boldsymbol{f}}(\boldsymbol{x}_{age}, \boldsymbol{x}_{year}) = \hat{\boldsymbol{\mu}}(\boldsymbol{x}_{age}, \boldsymbol{x}_{year})\hat{\boldsymbol{S}}(\boldsymbol{x}_{age}, \boldsymbol{x}_{year}).$$

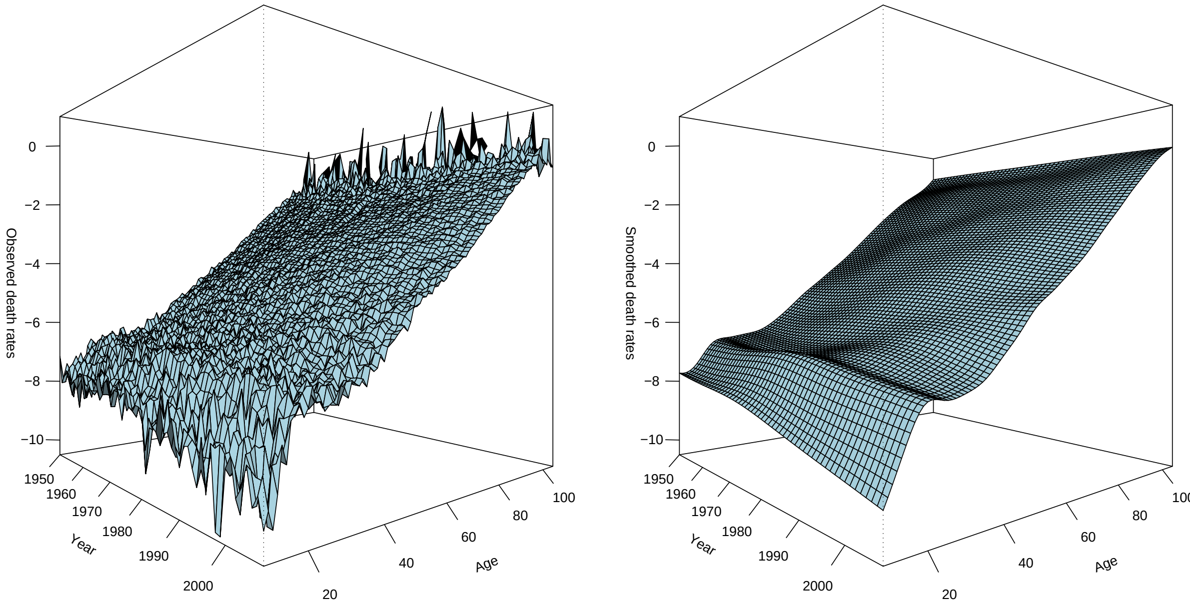


Figure 2: Observed death rates (left) and smoothed death rates using 2D Poisson P-spline smoothing (right), Danish males, 1950 to 2007. *Source*: HMD 2010

For example, figure 3 presents the smoothed density function for Danish males between 1950 and 2007. In order to follow and assess changes in $\hat{\boldsymbol{f}}$ over the years, summary measures such as $\hat{M}$ and $\hat{SD}(M+)$ for example can be computed. Since $\hat{\boldsymbol{f}}$ is a function of ages and years, both $\hat{M}$ and $\hat{SD}(M+)$ are functions and they respectively correspond to equations (2) and (3) below. Therefore, instead of providing discrete summary measures to monitor changes in the age-at-death distribution over time, the two dimensional smoothing approach leads to continuous summary measures.

$$\hat{M}(\boldsymbol{x}_{year}) = \max_{x_{year}} \hat{f}(\boldsymbol{x}_{age}, \boldsymbol{x}_{year}) \tag{2}$$

$$\widehat{SD(M+)}(\boldsymbol{x}_{year}) = \frac{\displaystyle\int_{\hat{M}}^{\infty} (x_{age} - \hat{M})^2 \hat{f}(x_{age}, \boldsymbol{x}_{year}) dx_{age}}{\displaystyle\int_{\hat{M}}^{\infty} \hat{f}(x_{age}, \boldsymbol{x}_{year}) dx_{age}} \tag{3}$$

Note that for a more complete monitoring of changes in variability in old age mortality, $\hat{SD}(M+)$ could be supplemented by a set of mode-based quantile functions. Indeed, mode-based quantiles provide greater flexibility than $\hat{SD}(M+)$ by allowing to focus on changes in variability which took place in any given neighbourhood above $\hat{M}$ (see Ouellette and Bourbeau (2009) for further details regarding mode-based quantiles).

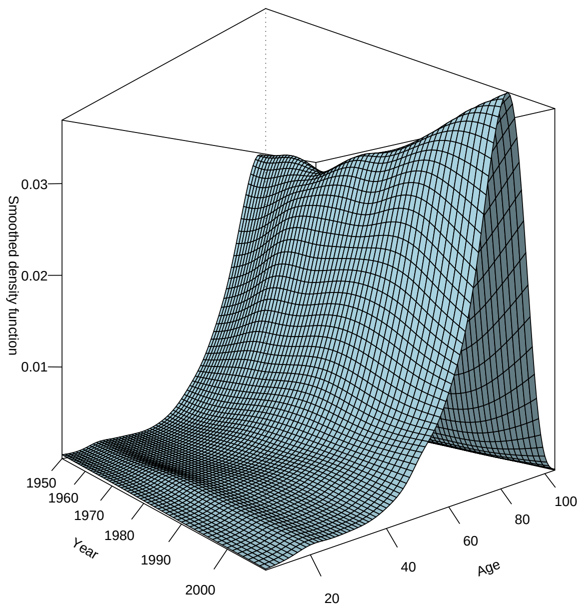


Figure 3: Two-dimensional density function describing the age-at-death distribution, Danish males, 1950 to 2007. *Source*: HMD 2010

Results and discussion

Figure 4 presents female changes in the estimated modal age at death $\hat{M}$ and standard deviation above the mode $\widehat{SD}(M+)$ in each of the ten countries. From the top panel of the figure, we see that females in most countries showed a sustained upward linear trend for $\hat{M}$ throughout the period under study. The average rate of increase for Japanese females has clearly been the highest of all, especially between 1960 and 2002 where it reached 3.3 months per year. However, an unexpected slowdown in $\hat{M}$ increasing trend occurred afterwards. The declining trend for US females from 2001 and onward was also unexpected. Consequently, the modal age at death for US females was estimated at 86.7 years in 2006, a level comparable to what they had achieved more than a decade earlier. In Denmark and the Netherlands, increases in $\hat{M}$ since 1950 have been less steady than in most of the other countries. Indeed, long pauses were recorded during the 1980s, 1990s, and even more recently among Danish females.

Based on the bottom panel of figure 4, we find that females of all countries displayed lower $\widehat{SD}(M+)$ at the end of the period studied compared to 1950. Thus, old-age compression of mortality has occurred since the beginning of the second half of the twentieth century among these females. However, for most countries, $\widehat{SD}(M+)$ did not decline steadily over the entire

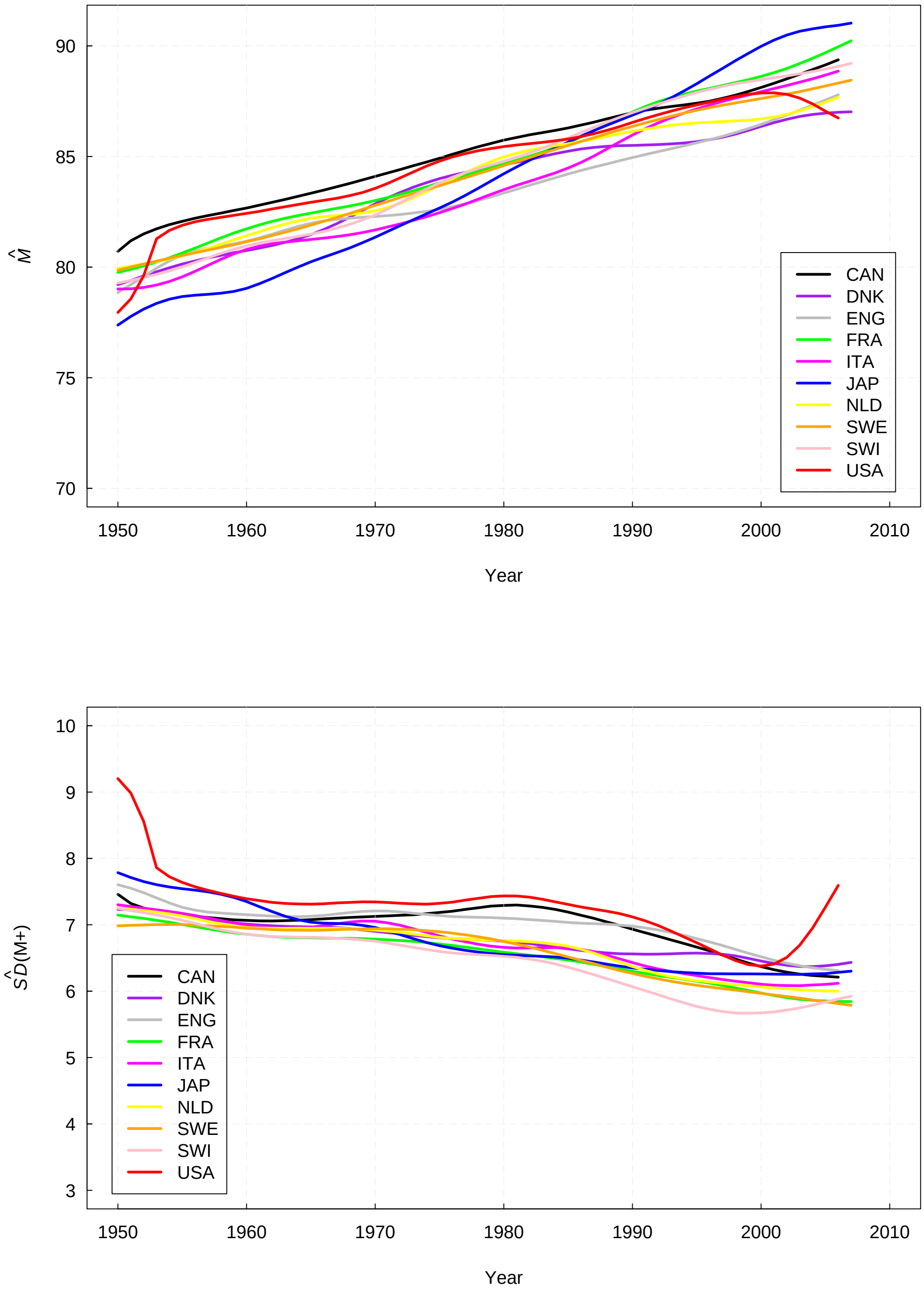
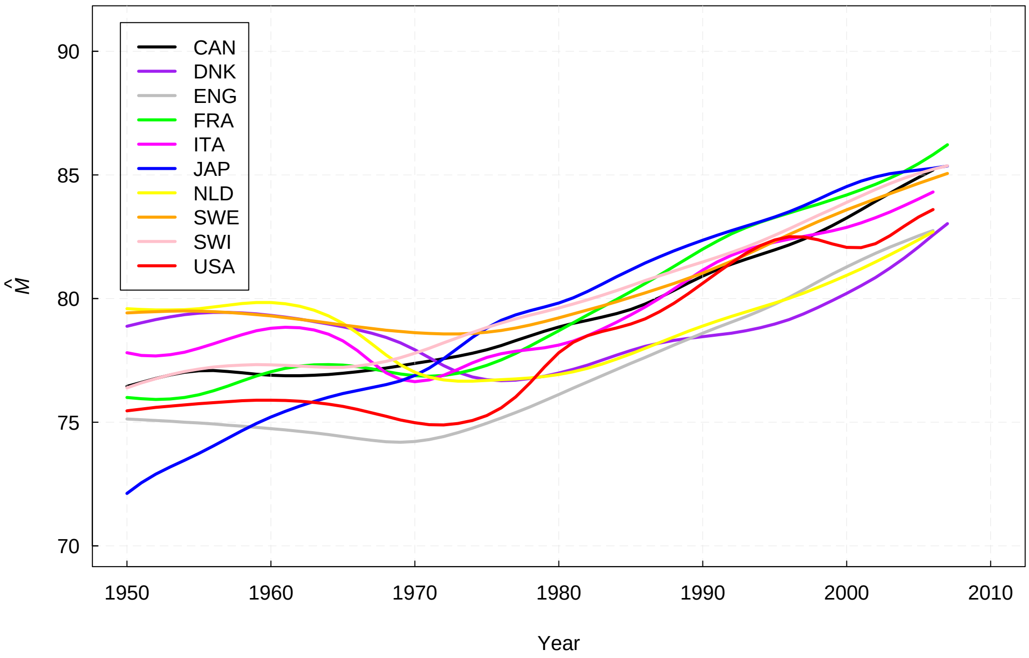


Figure 4: Estimated modal age at death (top) and standard deviation above it (bottom) using 2D Poisson P-spline smoothing since 1950, females. *Source*: HMD 2010

period. For example, $\widehat{SD}(M+)$ stagnated at about 7 years from 1950 up to the early 1970s among Swedish females. Furthermore, in the nine countries excluding Japan, the 1960s were years of very slow decline. Among Japanese females, the fall in $\widehat{SD}(M+)$ has levelled off since the early 1990s. Given that their $\hat{M}$ rose rapidly during most of that period, their distribution of ages at death at very old ages has been shifting to the right without changing its shape for several years now. Females from some of the other countries have also been thru such shifting episodes between 1950 and 2006 or 2007, but none has experienced one for as many successive years as the Japanese's.

Figure 5 shows male country-specific changes in the estimated modal age at death $\hat{M}$ and standard deviation above the mode $\widehat{SD}(M+)$. The top panel reveals that upward linear trends for $\hat{M}$ did not start much before the 1970s among males, except in Japan. Indeed, up to the 1970s, $\hat{M}$ rather stagnated or even decreased, probably because mortality reductions at ages older than $\hat{M}$ essential for its increase were very limited during those years (Canudas-Romo and Wilmoth, 2007). Afterwards however, the average rate of increase for males of all countries taken together was around 2.5 months per year and thus higher than among all females except those of Japan. The case of Japanese males stands out since $\hat{M}$ increased almost systematically over the whole period. Nevertheless, the rate of increase slowed down substantially among Japanese males during recent years. Thus, $\hat{M}$ for French males has been above that of Japanese males since 2005 and Swiss and Canadian males have now caught up with the Japanese's. Will increases in modal age at death resume in Japan? Given the long-term stability of historical increase in $\hat{M}$ in this country, it would be prudent to continue to foresee future increases among both sexes. In the USA, the upward trend since the mid-1970s for $\hat{M}$ has not been as steady as in the other countries. Indeed, $\hat{M}$ decreased abruptly in the late 1990s, but the increasing trend has resumed in the latest years.

If we focus on the bottom panel of figure 5, we see that $\widehat{SD}(M+)$ was not necessarily lower at the end of the period under study than in 1950. In other words, unlike what was found among females, compression above the mode did not occur in all ten countries over the period as a whole for males. Indeed, in Denmark, Italy and the Netherlands, $\widehat{SD}(M+)$ was around 7 years in both 1950 and 2006 or 2007. In between those years, $\widehat{SD}(M+)$ rose and then fell, thus revealing successive episodes of old-age decompression and compression of mortality. However, if we limit ourselves to the period starting with the onset of $\hat{M}$ increase for each country (around the 1970s, except Japan), it coincides with episodes of old-age compression of mortality because $\widehat{SD}(M+)$ was falling. Among Japanese males, $\widehat{SD}(M+)$ has been stagnating since the early 1990s and this has been paralleled by increases in $\hat{M}$. Thus, their distribution of ages at death at very old ages has shifted to the right, keeping an intact shape.



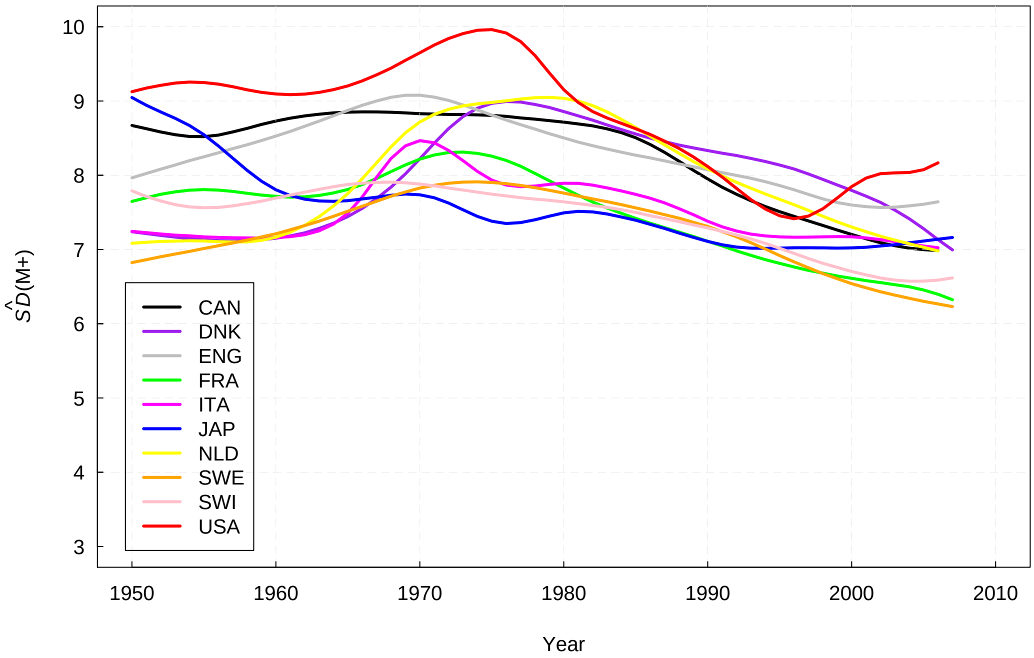


Figure 5: Estimated modal age at death (top) and standard deviation above it (bottom) using 2D Poisson P-spline smoothing since 1950, males. *Source*: HMD 2010

Overall, the results described above are consistent with recent findings on changes in the age-at-death distribution at advanced ages (Canudas-Romo, 2008; Cheung et al., 2005, 2008, 2009; Cheung and Robine, 2007; Ouellette and Bourbeau, 2009; Thatcher et al., 2010). The two-dimensional nonparametric method presented in this paper is a generalization of previous work by Ouellette and Bourbeau (2009) and it does not make any rigid assumption about the functional form of mortality rates along ages and years. It solely rests on the assumption that mortality changes over ages and years are regular, and that the erratic behaviour is essentially explained by the randomness of rates (see figure 2). For countries with reliable data, this flexible smoothing approach represents a natural choice, notably for monitoring changes that have occurred at older ages over time. Furthermore, the fact that the model takes simultaneously into account the continuous nature of mortality progression over ages and time can be particularly useful for countries with small population size or for analysing population subgroups.

Acknowledgements

This project was funded by the Social Sciences and Humanities Research Council of Canada. Part of the research was conducted while the first author visited the Statistical Demography Laboratory at the Max Planck Institute for Demographic Research. The authors are grateful to Jutta Gampe and Carlo G. Camarda for their helpful suggestions and support.

References

Bongaarts, J. (2005). Long range trends in adult mortality: models and projection methods. *Demography 42*(1), 23–49.

Brouhns, N., M. Denuit, and J. K. Vermunt (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics & Economics 31*, 373–393.

Camarda, C. G. (2008). *Smoothing Methods for the Analysis of Mortality Development*. Ph. D. thesis, Universidad Carlos III de Madrid and Max Planck Institute for Demographic Research.

Camarda, C. G. (2009). *MortalitySmooth: Smoothing Poisson counts with P-splines.* R package version 1.0.

Canudas-Romo, V. (2008). The modal age at death and the shifting mortality hypothesis. *Demographic Research 19*, 1179–1204.

Canudas-Romo, V. and J. R. Wilmoth (2007). Record measures of longevity. Paper presented at the Annual Meeting of the Population Association of America, March 29–31 2007.

Cheung, S. L. K. and J.-M. Robine (2007). Increase in common longevity and the compression of mortality: the case of Japan. *Population Studies 61*(1), 85–97.

Cheung, S. L. K., J.-M. Robine, and G. Caselli (2008). The use of cohort and period data to explore changes in adult longevity in low mortality countries. *Genus LXIV*(1-2), 101–129.

Cheung, S. L. K., J.-M. Robine, F. Paccaud, and A. Marazzi (2009). Dissecting the compression of mortality in Switzerland, 1876-2005. *Demographic Research 21*, 569–598.

Cheung, S. L. K., J.-M. Robine, E. J.-C. Tu, and G. Caselli (2005). Three dimensions of the survival curve: horizontalization, verticalization, and longevity extension. *Demography 42*(2), 243–258.

Clayton, D. and E. Schifflers (1987). Models for temporal variation in cancer rates. ii: Age-period-cohort models. *Statistics in Medicine 6*, 469–481.

Currie, I. D., M. Durban, and P. H. C. Eilers (2004). Smoothing and forecasting mortality rates. *Statistical Modelling 4*, 279–298.

Currie, I. D., M. Durban, and P. H. C. Eilers (2006). Generalized linear array models with applications to multidimensional smoothing. *Journal of the Royal Statistical Society 68*, 259–280.

Eilers, P. H. C., I. D. Currie, and M. Durban (2006). Fast and compact smoothing on large multidimensional grids. *Computational Statistics and Data Analysis 50*, 61–76.

Eilers, P. H. C. and B. D. Marx (1996). Flexible smoothing with B-splines and penalties (with discussion). *Statistical Science 11*(2), 89–102.

Eilers, P. H. C. and B. D. Marx (2002). Multivariate calibration with temperature interaction using two-dimensional penalized signal regression. *Chermometrics and Intelligent Laboratory Systems 66*, 159–174.

Human Mortality Database (2010). University of California, Berkley (USA) and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org (data downloaded on [01/03/2010]).

Jeune, B. and J. W. Vaupel (1995). *Exceptional Longevity: From Prehistory to the Present.* Odense: Odense University Press.

Kannisto, V. (1996). *The Advancing Frontier of Survival Life Tables for Old Age*, Monographs on Population Aging, 3. Odense: Odense University Press.

Kannisto, V. (2000). Measuring the compression of mortality. *Demographic Research 3*(6).

Kannisto, V. (2001). Mode and dispersion of the length of life. *Population: An English Selection 13*(1), 159–171.

Kannisto, V., J. Lauristen, A. R. Thatcher, and J. W. Vaupel (1994). Reductions in mortality at advanced ages: several decades of evidence from 27 countries. *Population and Development Review 20*(4), 793–809.

Lee, R. C. and L. R. Carter (1992). Modeling and forecasting u.s. mortality. *Journal of the American Statistical Association 87*, 659–675.

Lexis, W. (1877). *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft [On the Theory of Mass Phenomena in Human Society].* Freiburg i.B.: Fr. Wagner'sche Buchhandlung.

Lexis, W. (1878). Sur la durée normale de la vie humaine et sur la théorie de la stabilité des rapports statistiques [On the normal human lifespan and the theory of the stability of statistical ratios]. *Annales de démographie internationale 2*, 447–460.

Ouellette, N. and R. Bourbeau (2009). Changes in the age-at-death distribution in low mortality countries: A nonparametric approach. Presented at the XXVIe International Population Conference of the IUSSP held in Marrakech (Morocco).

Thatcher, A. R., S. L. K. Cheung, S. Horiuchi, and J.-M. Robine (2010). The compression above the mode. *Demographic Research 22*, 505–538.

Wilmoth, J. R. and S. Horiuchi (1999). Rectangularization revisited: Variability of age at death within human populations. *Demography 36*(4), 475–495.

Yashin, A. L., A. S. Begun, S. I. Boiko, S. V. Ukrainseva, and J. Oeppen (2001). The new trends in survival improvement require a revision of traditional gerontological concepts. *Experimental Gerontology 37*(1), 157–167.

Appendix

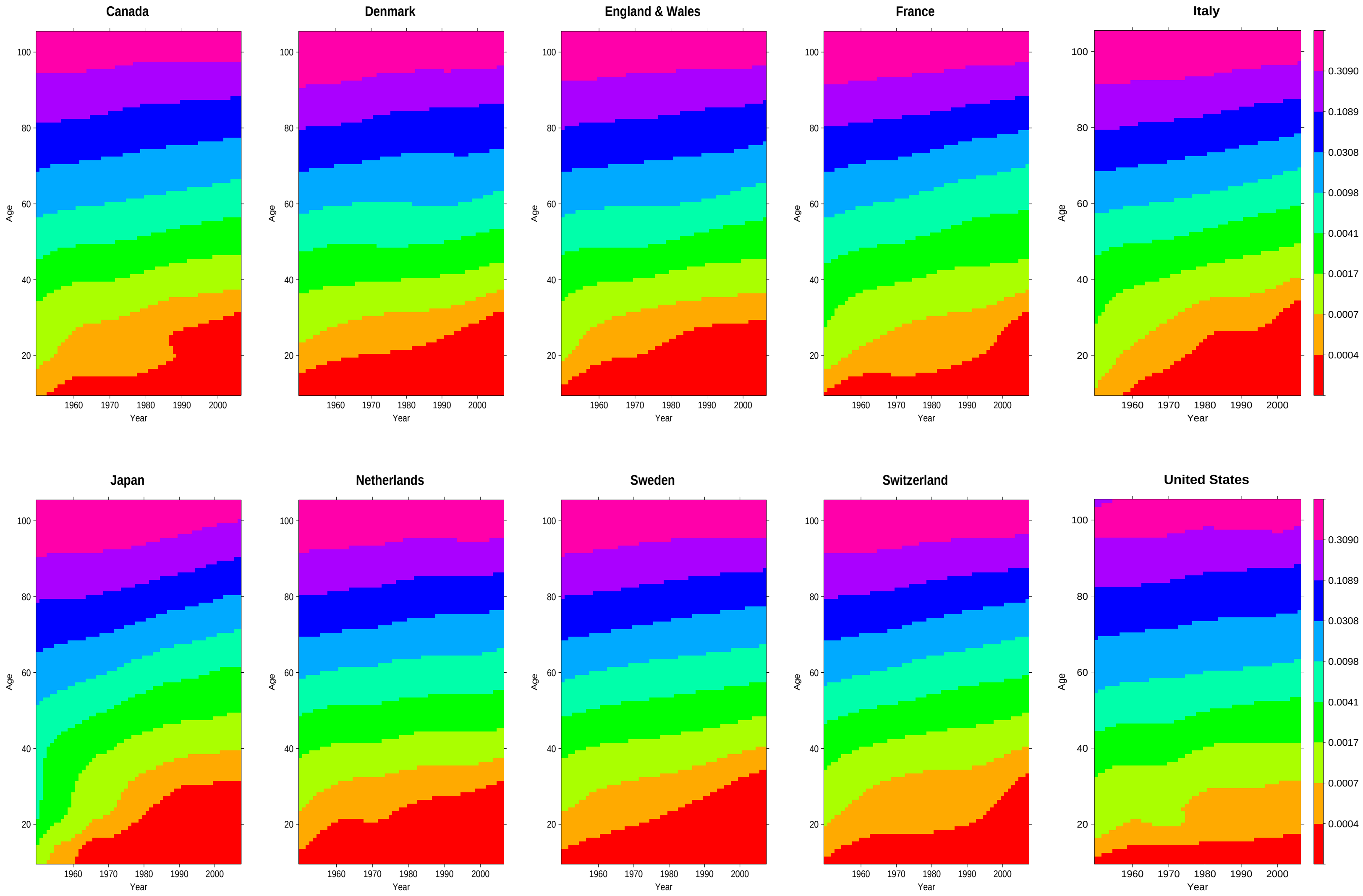


Figure 6: Smoothed mortality surfaces among females since 1950. *Source*: HMD 2010

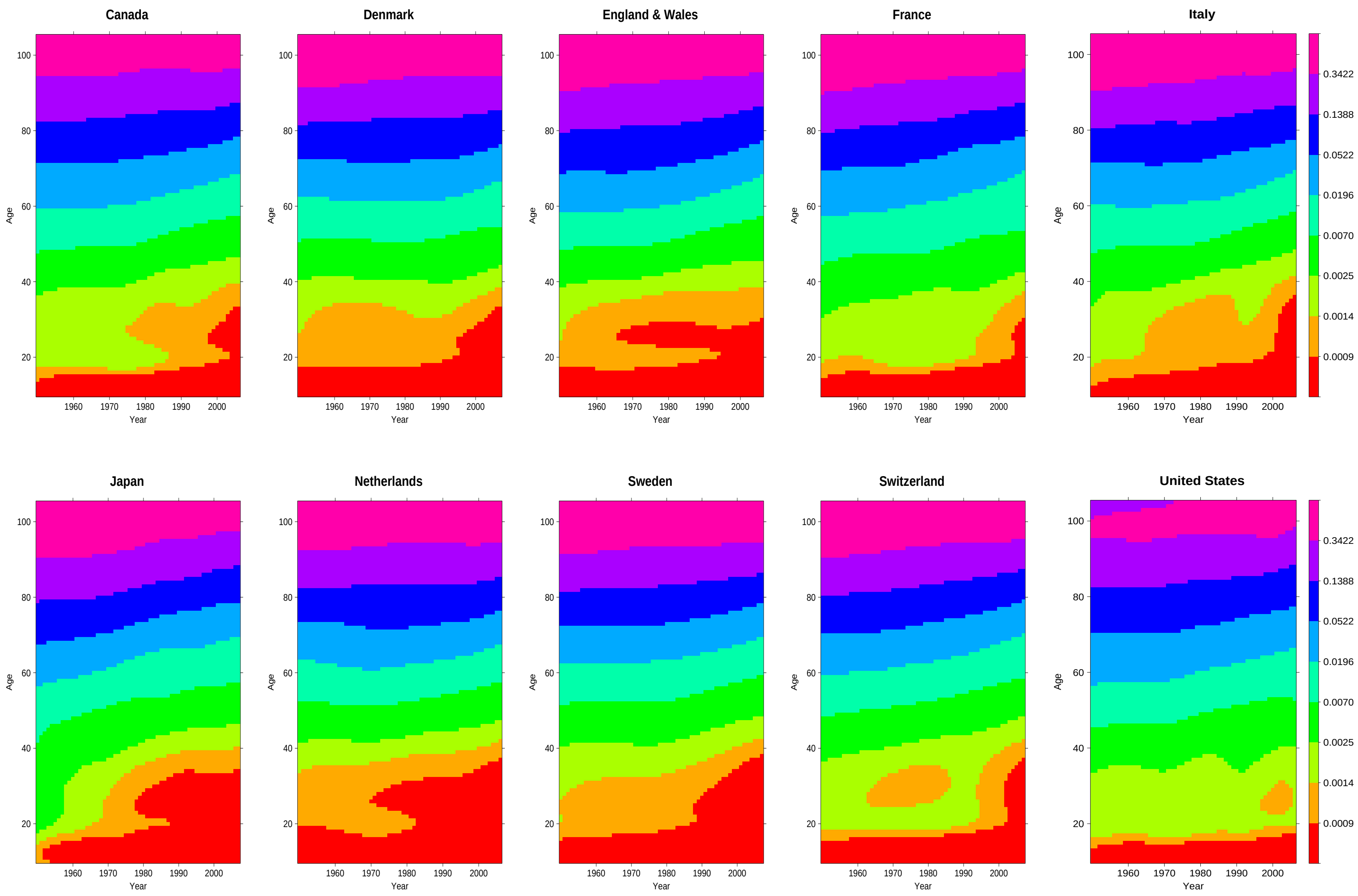


Figure 7: Smoothed mortality surfaces among males since 1950. *Source*: HMD 2010