

Paper accepted for presentation at  
The 2011 Annual Meeting of the Population Association of America

March 31-April 2, 2011  
Washington, DC

*Session 146 "Becoming a Centenarian"*

## **Old-Age Mortality Measurement and Modeling**

**Natalia S. Gavrilova, Leonid A. Gavrilov,**

Center on Aging, NORC/University of Chicago, 1155 East 60<sup>th</sup> Street,  
Chicago, IL 60637

Address for correspondence:  
Dr. Natalia S. Gavrilova, Center on Aging  
NORC/University of Chicago  
1155 East 60<sup>th</sup> Street, Chicago, IL 60637  
Fax: (773) 256-6313; Phone: (773) 702-1375  
E-mail: [gavrilova@longevity-science.org](mailto:gavrilova@longevity-science.org)

## **Abstract**

A growing number of persons living beyond age 100 emphasize the need for accurate measurement and modeling of mortality at advanced ages. However, mortality measurement at extremely old ages poses serious challenges to researchers: age reporting at extreme ages may be inaccurate; small number of persons at advanced ages requires pooling together data for people belonging to different birth cohorts, which results in data heterogeneity; finally, standard approach based on annual mortality estimates may not be appropriate for studies of mortality at extreme old ages when death rates are particularly high.

The study uses data from the Social Security Administration Death Master File (DMF) to measure mortality in more homogeneous single-year birth cohorts and in narrow (monthly) age intervals. Mortality after age 88 was estimated for 1881-1895 single-year extinct birth cohorts. Data quality control checks demonstrated that deceleration of mortality in later life is more expressed for data of lower quality. Additional quality checks using female-to-male ratio at advanced ages identified the age 107 years as the upper limit for acceptable quality of age reporting. Model comparison using Bayesian information criterion (BIC) found that the Gompertz model provides better fitting of hazard rates than the logistic model in the age interval 88-106 years and for data with reasonably good quality. Hazard rate estimates at old ages obtained on the basis of DMF data agree well with estimates calculated on the basis of 1900 actuarial cohort life table.

Study of several single-year extinct U.S. birth cohorts found that mortality trajectory at advanced ages follows the Gompertz law up to the ages 106 years without significant mortality deceleration. Earlier reports of mortality deceleration for ages below 106 years appears to be a result of age exaggeration, use of improper estimates of hazard rate and perhaps data heterogeneity. The phenomenon of mortality deceleration at older ages vanishes for datasets of higher quality. At the same time, our findings do not exclude possibility of mortality deceleration at later ages (after age 107 years).

## 1. Introduction

Accurate estimates of mortality at advanced ages are essential for improving forecasts of mortality and predicting the population size of the oldest old age group. It is believed that mortality at advanced ages has a tendency to deviate from the Gompertz law (Gavrilov and Gavrilova 1991), so that the logistic model is suggested for fitting human mortality after age 85 years (Horiuchi and Wilmoth 1998). The estimates of mortality force at extreme ages are difficult because of small numbers of survivors to these ages in most countries. Data for extremely long-lived individuals are scarce and subjected to age exaggeration. Traditional demographic estimates of mortality based on period data encounter well known denominator problem. More accurate estimates are obtained using the method of extinct generations (Vincent 1951). In order to minimize statistical noise in estimates of mortality at advanced ages researches are forced to pool data for several calendar periods (Depoid 1973; Thatcher 1999). Single-year life tables for many countries have very small numbers of survivors to age 100 that makes estimates of mortality at advanced ages unreliable. On the other hand, aggregation of deaths for several calendar periods creates a heterogeneous mixture of cases from different birth cohorts. Theoretical models suggest that mortality deceleration at advanced ages may be caused by population heterogeneity even if individual risk of death follows the Gompertz law (Beard 1959; Beard 1971; Vaupel, Manton and Stallard 1979). In addition to that, many standard assumptions used in hazard rate estimation are not valid for extreme old ages when mortality is particularly high.

Thus the estimation of hazard rate at extremely old ages poses several serious challenges to researchers:

- (1) The observed mortality deceleration may be at least partially related to mixing different birth cohorts with different mortality (heterogeneity effect);
- (2) Standard assumptions of hazard rate estimates may be invalid when risk of death is extremely high at old ages;
- (3) Ages of very old people may be exaggerated.

One way of obtaining estimates of mortality at extreme ages is to pool together international records of persons surviving to extreme ages with subsequent efforts of strict age validation (Robine et al. 2005; Robine and Vaupel 2001). This approach helps to resolve the third problem mentioned above but does not allow researchers to resolve the first two problems because of inevitable data heterogeneity when data for people belonging to different birth cohorts and countries are pooled together.

In this paper we propose an alternative approach, which allows us to resolve partially these problems. This approach is based on using data from the Social Security Administration Death Master File (DMF), which allows compiling data for large single-year birth cohorts. Some already extinct birth cohorts covered by DMF could be studied by the method of extinct generations. Availability of month of birth and month of death information provides a unique opportunity to obtain hazard rate estimates for every month of age. Possible ways of resolving the age exaggeration problem in hazard rate estimation are also elaborated.

## 2. Mortality at Advanced Ages: A Historical Review

The history of mortality studies at extreme ages is very rich in ideas and findings. Early studies starting with Gompertz (1825) himself suggested that the Gompertz law of

mortality is not applicable to extreme old ages, and that mortality deceleration and leveling-off takes place at advanced ages (for a thoughtful historical review of studies on mortality deceleration at extreme old ages, see Olshansky 1998). In 1939 two British researchers, Greenwood and Irwin, published an article “Biostatistics of Senility,” with an intriguing finding that mortality force stops increasing with age at extreme old ages and becomes constant (Greenwood and Irwin, 1939). This study, lead by the famous British statistician and epidemiologist, Major Greenwood, may be interesting to discuss here because it describes the mortality pattern at advanced ages for humans reported by researchers many years later (Depoid 1973; Gavrilov 1984; Horiuchi and Wilmoth 1998; Thatcher 1999; Thatcher, Kannisto and Vaupel 1998).

The first important finding was formulated by Greenwood and Irwin in the following way: “...*the increase of mortality rate with age advances at a slackening rate, that nearly all, perhaps all, methods of graduation of the type of Gompertz’s formula over-state senile mortality*” (Greenwood and Irwin, 1939, p. 14). This observation was reported later by many authors (Depoid 1973; Gavrilov and Gavrilova 1991; Horiuchi and Wilmoth 1998; Thatcher 1999; Thatcher et al. 1998), and it is now known as the “late-life mortality deceleration.”

The authors also suggested “*the possibility that with advancing age the rate of mortality asymptotes to a finite value*” (Greenwood and Irwin, 1939, p. 14). Their conclusion that mortality at exceptionally high ages follows a first order kinetics (also known as the law of radioactive decay with exponential decline in survival probabilities) was supported later by other researchers, including A.C. Economos (Economos 1979; Economos 1980), who reported that this law holds for humans and laboratory animals (linear decrease for the logarithm of the number of survivors). This observation is known now as the “mortality leveling-off” at advanced ages, and as the “late-life mortality plateau” (Curtsinger, Gavrilova and Gavrilov 2006).

Moreover, Greenwood and Irwin made the first estimates for the asymptotic value of human mortality (one-year probability of death,  $q_x$ ) at extreme ages using data from the life insurance company. According to their estimates, “... *the limiting values of  $q_x$  are 0.439 for women and 0.544 for men*” (Greenwood and Irwin, 1939, p. 21). It is interesting that these first estimates are very close to estimates obtained later using more numerous and accurate human data including recent data on supercentenarians (Robine and Vaupel, 2001). The authors also suggested an explanation of this phenomenon. According to Greenwood and Irwin (1939), centenarians live in a more protected environment than younger age groups and hence have lower risk of death than predicted by the Gompertz formula.

The actuaries including Gompertz himself were the first who noticed this phenomenon of mortality deceleration. They also proposed a logistic formula for fitting mortality growth with age in order to account for mortality leveling-off at advanced ages (Perks, 1932; Beard, 1959, 1971). British actuary, Robert Eric Beard, introduced a model of population heterogeneity with gamma distributed individual risk in order to explain mortality deceleration at older ages (Beard 1959). This explanation is now considered to be the most common explanation of mortality deceleration phenomenon (Horiuchi and Wilmoth 1998).

The same phenomenon of “almost non-aging” survival dynamics at extreme old ages is detected in many other biological species. In some species a mortality plateau can occupy a sizable part of their life (Carey et al. 1992; Gavrilov and Gavrilova 2006). Biologists were aware of mortality leveling-off since the 1960s. For example, Lindop (Lindop 1961) applied Perks formula in order to account for mortality deceleration at older ages in mice. George Sacher believed that the observed mortality deceleration in

mice and rats can be explained by population heterogeneity: “one effect of such residual heterogeneity is to bring about a decreased slope of the Gompertzian at advanced ages. This occurs because sub-populations with the higher injury levels die out more rapidly, resulting in progressive selection for vigour in the surviving populations” (Sacher, 1966, p. 435). Strehler and Mildvan considered mortality deceleration at advanced ages as a prerequisite for all mathematical models of aging to explain (Strehler and Mildvan 1960).

Later Economos published a series of articles claiming a priority in the discovery of a “non-Gompertzian paradigm of mortality” (Economos 1979; Economos 1980, 1983, 1985). He found that mortality leveling-off is observed in rodents (guinea pigs, rats, mice) and invertebrates (nematodes, shrimps, bdelloid rotifers, fruit flies, degenerate medusae *Campanularia Flexuosa*). In the 1990s, the phenomenon of mortality deceleration and leveling-off became widely known after publications, which demonstrated mortality leveling-off in large samples of *Drosophila melanogaster* (Curtsinger et al. 1992; Curtsinger et al. 2006) and medflies *Ceratitis capitata* (Carey et al. 1992). Mortality plateaus at advanced ages are observed for some other insects: house fly *Musca vicina*, blowfly *Calliphora erythrocephala* (Gavrilov 1980), house fly *Musca domestica* (Gavrilov and Gavrilova 2006), fruit flies *Anastrepha ludens*, *Anastrepha obliqua*, *Anastrepha serpentine*, parasitoid wasp *Diachasmimorpha longicaudtis* (Vaupel et al. 1998), and bruchid beetle *Callosobruchus maculatus* (Tatar, Carey and Vaupel 1993). Interestingly, the failure kinetics of manufactured products (steel samples, industrial relays, and motor heat insulators) also demonstrates the same “non-aging” pattern at the end of their “lifespan” (Economos 1979).

Population heterogeneity, first proposed by Beard in 1959, is by far the most common explanation of mortality deceleration (Horiuchi and Wilmoth 1998). Another explanation of this phenomenon comes from the reliability theory of aging, which explains mortality leveling-off by an exhaustion of organism’s redundancy (reserves) at extremely old ages so that every random hit results in death (Gavrilov and Gavrilova 1991; Gavrilov and Gavrilova 2001; Gavrilov and Gavrilova 2006). There is also an opinion that lower risks of death for older people are due to their less risky behavior (Greenwood and Irwin 1939). Finally, some researchers suggest evolutionary explanations for the phenomenon of mortality leveling-off (Charlesworth 2001; Mueller and Rose 1996).

The existence of mortality plateaus is well established for a number of lower organisms, mostly insects. In the case of mammals, data are much more controversial. Lindop (1961) and Sacher (1966) reported short-term periods of mortality deceleration in mice at advanced ages and even used Perks formula in their analyses. However, Austad later argued that rodents do not demonstrate mortality deceleration even in the case of large samples (Austad 2001). Study of baboons found no mortality deceleration at advanced ages (Bronikowski et al. 2002). In the case of humans this problem is not yet resolved, because of scarceness of data and/or their low reliability. Thus, more studies on larger human birth cohorts are required to establish with certainty the true mortality trajectory at advanced ages.

### 3. Hazard rate (mortality force) estimation at advanced ages

A conventional way to obtain estimates of mortality at advanced ages is a construction of demographic life table with probability of death ( $q_x$ ) as one of important life table functions. Although probability of death is a useful indicator for mortality studies, it may be not the most convenient one for studies of mortality at advanced ages. First, the values of  $q_x$  depend on the length of the age interval  $\Delta x$  for which it is calculated. This

hampers both analyses and interpretation. For example, if one-year probability of death follows the Gompertz law of mortality, probability of death calculated for other age interval does not follow this law (Gavrilov and Gavrilova 1991; Le Bras 1976). Thus it turns out that the shape of age trajectory for  $q_x$  depends on the arbitrary choice of age interval. Also, by definition  $q_x$  is bounded by unity, which would inevitably produce apparent mortality deceleration when death rates are particularly high.

More useful indicator of mortality at advanced age is instantaneous mortality rate (mortality force) or hazard rate,  $\mu_x$  which is defined as follows:

$$\mu_x = -\frac{dN_x}{N_x dx} = -\frac{d \ln(N_x)}{dx} \approx -\frac{\Delta \ln(N_x)}{\Delta x}$$

where  $N_x$  is a number of living individuals exposed to risk of death at age  $x$ . It follows from the definition of hazard rate that it is equal to the rate of decrease of logarithmic survival function with age. In actuarial practice, hazard rate is often called mortality force as it was done in the original paper by Benjamin Gompertz (Gompertz 1825). Hazard rate does not depend on the length of the age interval (it is measured at the instant of time  $x$ ), has no upper boundary and has a dimension of rate ( $\text{time}^{-1}$ ). It should also be noted that the famous law of mortality, the Gompertz law, was first proposed for fitting the age-specific hazard rate function rather than probability of death (Gompertz 1825).

The empirical estimates of hazard rates are often based on suggestion that age-specific mortality rate or death rate (number of deaths divided by exposure) is a good estimate of theoretical hazard rate. One of the first empirical estimates of hazard rate was proposed by George Sacher (Sacher 1956; Sacher 1966):

$$\mu_x = \frac{1}{\Delta x} [\ln(l_{x - \frac{\Delta x}{2}}) - \ln(l_{x + \frac{\Delta x}{2}})] = \frac{1}{\Delta x} \ln \left( \frac{l_{x - \frac{\Delta x}{2}}}{l_{x + \frac{\Delta x}{2}}} \right)$$

This estimate is unbiased for slow changes in hazard rate if  $\Delta x \Delta \mu_x \ll 1$  (Sacher, 1966) and was shown to be the maximum likelihood estimate (Gehan and Siddiqui 1973). A simplified version of Sacher estimate (for small age intervals equal to unity) is often used in biodemographic studies of mortality:  $\mu_x = -\ln(1-q_x)$ . This estimate was initially suggested by Gehan who called it a 'Sacher' estimate (Gehan 1969; Gehan and Siddiqui 1973). It is based on the assumption that hazard rate is constant over age interval and is shifted by one half of a year to younger ages compared to the original Sacher estimate.

Another empirical estimate of hazard rate, often used in life table construction (Klein and Moesberger 1997), is the actuarial estimate, which is calculated in the following way (Kimball 1960):

$$\mu_x = \frac{2q_x}{\Delta x (2 - q_x)} = \frac{2}{\Delta x} \frac{l_{x - \Delta x} - l_x}{l_{x - \Delta x} + l_x}$$

This estimate assumes uniform distribution of deaths over the age interval and is bounded by  $2/\Delta x$ , so this is not the best estimate of hazard rate at extreme old ages when death rates are particularly high (Gavrilov and Gavrilova 1991).

At advanced ages, when death rates are very high, the assumptions about small changes in hazard rate or a constant hazard rate within the age interval become questionable. The same is true for the assumption of uniform distribution of deaths within the age interval. Simulation studies showed that the bias in hazard rate estimation increases with the increase of the age interval (Kimball 1960). Narrowing the age interval for hazard rate estimation from one year to one month helps to improve the accuracy of hazard rate estimation.

#### **4. Social Security Administration Death Master File as a source of mortality data for advanced ages.**

Social Security Administration Death Master File (DMF) is a publicly available data source that allows a search for deceased individuals in the United States using various search criteria: birth date, death date, first and last names, social security number, place of last residence, etc. This resource covers deaths that occurred in the period 1937-2010 (Faig 2001) and captures about 95% of deaths recorded by the National Death Index (Sesso, Paffenbarger and Lee 2000). According to other estimates, DMF covers about 90 percent of all deaths for which death certificates are issued (see Faig, 2001) and about 92-96 percent of deaths for persons older than 65 years (Hill, Rosenwaike, 2001).

Social Security Administration Death Master File (DMF) was used in our study of age-related mortality dynamics after ages 85 years. The advantage of this data source is that some already extinct birth cohorts covered by DMF could be studied by the method of extinct generations (Kannisto 1988, 1994; Vincent 1951). Information available in DMF includes: names of the deceased, his/her social security number, date, month, year of birth, month and year of death, state of SSN issuance, place of the last residence.

In this study information from the DMF was collected for individuals who lived 88 years and over and died before 2011. DMF database is unique because it represents mortality experience for very large birth cohorts of the oldest-old persons. In this study mortality measurements were made for cohorts, which are more homogeneous in respect to the year of birth and historical life course experiences. Availability of month of birth and month of death information provides a unique opportunity to obtain hazard rate estimates for every month of age, which is important given extremely high mortality after age 100 years. Despite certain limitations, this data source allows researchers to obtain detailed estimates of mortality at advanced ages. We already used this data resource for centenarians' age validation in the study of centenarian family histories (Gavrilova and Gavrilov 2007). This data resource is also useful in mortality estimates for several extinct or almost extinct birth cohorts in the United States.

#### **5. Hazard rate estimates at advanced ages using data from the Social Security Death Master File**

In this study we collected information from the DMF publicly available at Rootsweb.com website. The total number of records collected from this resource was about 9 million with more than 900,000 records belonging to persons who lived 100 years and over. We obtained data for persons who died before 2011 and were born in 1875-1895. Persons born in these years and alive in 2010 should survive to at least 115

years, which can be considered as very unlikely event. Thus, the 1875-1895 birth cohorts in this sample may be considered as practically extinct. Assuming that the number of living persons belonging to these birth cohorts in 2010 is close to zero, it is possible to construct a cohort life table using well-known method of extinct generations, which was suggested and explained by Vincent (1951) and developed further by Kannisto (Kannisto 1994).

DMF provides information about years and months of birth and death. However information on exact day of death is not available for all records, so we were not able to calculate lifespan in days. Nevertheless we were able to calculate individual lifespan with one-month accuracy, which is still higher than the accuracy of traditional yearly estimate of lifespan. In the first stage of our analyses we calculated an individual life span in completed months:

$$\text{Lifespan in months} = (\text{death year} - \text{birth year}) \times 12 + \text{death month} - \text{birth month}$$

Having this information it is possible to estimate hazard rates at each month of age by standard methods of survival analysis. All calculations were done using the Stata statistical software, release 11 (StataCorp 2009). This software calculates nonparametric estimates of major survival functions including the Nelson-Aalen estimator of hazard rate (force of mortality). Note that hazard rate,  $\mu(x)$ , in contrast to probability of death,  $q(x)$ , has a dimension of time frequency, because of the time interval in the denominator (reciprocal time,  $\text{time}^{-1}$ ). Thus the values of hazard rates depend on the chosen units of time measurement ( $\text{day}^{-1}$ ,  $\text{month}^{-1}$  or  $\text{year}^{-1}$ ). In this study, survival times were measured in months, so the estimates of hazard rates initially had a dimension of  $\text{month}^{-1}$ . For the purpose of comparability with other published studies, which typically use the  $\text{year}^{-1}$  time scale, we transformed the monthly hazard rates to the more conventional units of  $\text{year}^{-1}$ , by multiplying these estimates by a factor of 12 (one month in the denominator of hazard rate formula is equal to 1/12 year). It should be noted that hazard rate, in contrast to probability of death, can be greater than 1, and therefore its logarithm can be greater than 0 (and we indeed observed these values at extreme old ages in some rare cases as will be described later). In this paper we focus our analyses on 1881-1895 birth cohorts, because we found that data quality for cohorts born before 1881 is not particularly good.

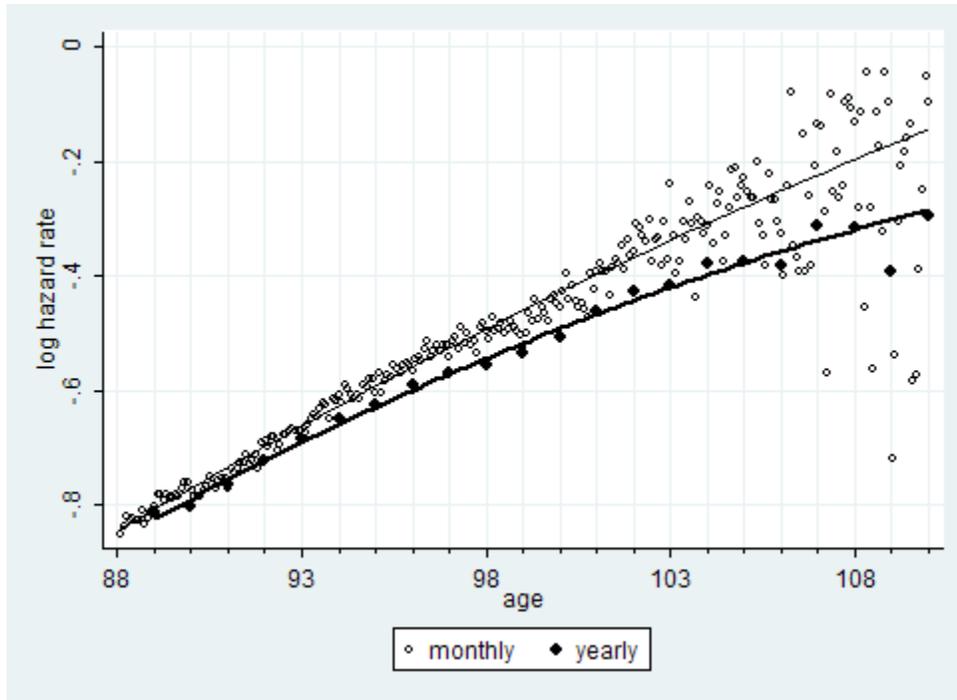
In our study we used Nelson-Aalen hazard rate estimates provided by Stata statistical software (StataCorp 2009). In fact, the Nelson-Aalen method was initially proposed for cumulative hazard rate estimation (particularly for right-censored survival data) (Klein and Moesberger 1997). In Stata hazard rate estimation is made by taking the steps of the Nelson-Aalen cumulative hazard function (Cleves et al. 2008), so that for each observed time of death,  $x_j$  the estimated hazard contribution is:

$$\Delta \hat{H}(x_j) = \hat{H}(x_j) - \hat{H}(x_{j-1})$$

where  $\hat{H}(x)$  is an estimate of cumulative hazard function.

The way of hazard rate estimation conducted in Stata is similar to calculation of life table probability of death (StataCorp 2009), i.e. the number of deaths in the studied age interval is divided by the number alive at the beginning of age interval. At advanced ages when mortality is high and for relatively large age intervals, number of persons exposed to risk of death in the middle of age interval is substantially lower than the number alive at the beginning of age interval. This would result in downward bias in hazard rate estimates at advanced ages, which we observed in our study when the

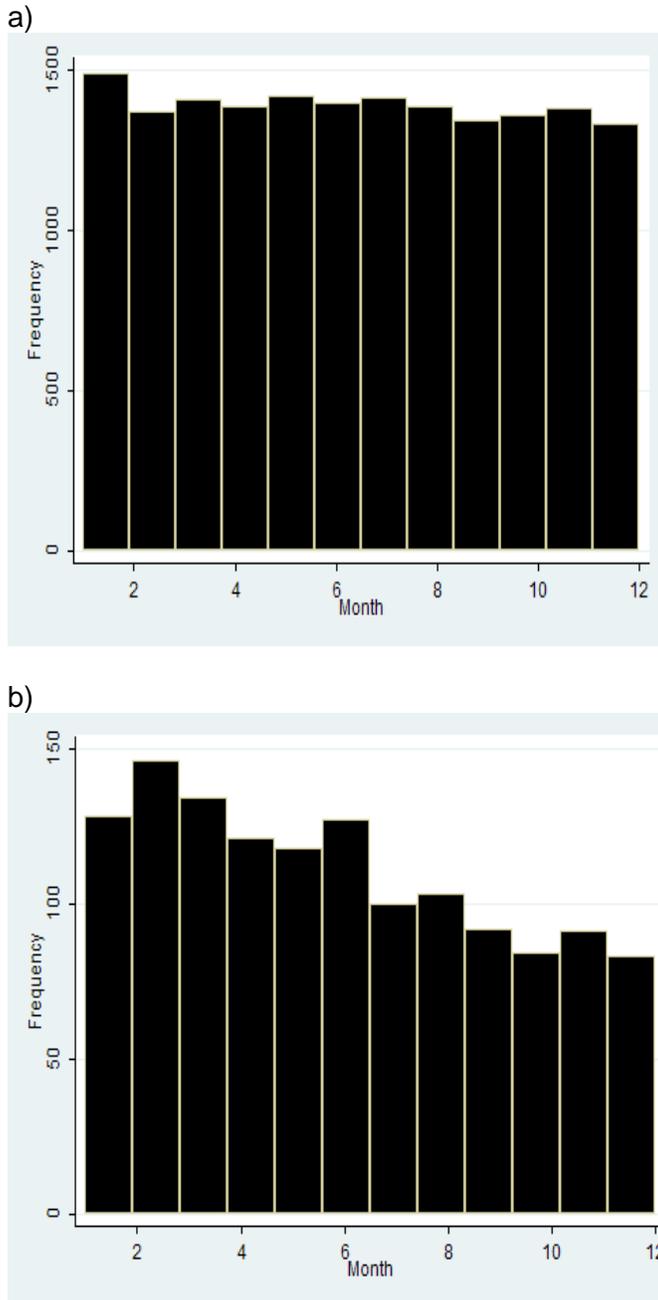
Nelson-Aalen estimates were applied to yearly age intervals (Figure 1, lower line). However, for smaller monthly age intervals, the problem described above is not so crucial and the Nelson-Aalen method still can be applied (see Figure 1, upper line).



**Figure 1. Logarithm of hazard rate (per year) as a function of age calculated for monthly and yearly age intervals. Hazard rate estimates were obtained using Nelson-Aalen estimate. Solid lines show polynomial fit with quadratic term. Data for 1891 birth cohort from the Social Security Administration Death Master File.**

Note that hazard rate values calculated on monthly basis demonstrate steady exponential growth up to the age 105 years (Figure 1). After the age 105 years mortality trajectories show a small tendency for deceleration with age. Study of hazard rate estimates for several other single-year birth cohorts (1886-1895) found a similar age pattern for mortality at advanced ages.

Another issue that makes monthly age intervals preferable for hazard rate estimation is the pattern of deaths distribution within studied age interval at extremely old ages. This issue is important because some estimates of hazard rate (for example, the actuarial estimate of hazard rate) are based on the assumption of approximately uniform distribution of deaths within one-year age intervals (Gehan and Siddiqui 1973; Kimball 1960; Watson and Leadbetter 1964). Indeed this assumption is consistent with the data for 90-year olds, presented in Figure 2a. However for 102-year olds, the distribution of deaths is obviously not uniform, with far more deaths occurring at the beginning of age interval (Figure 2b). To alleviate this problem, it is preferable to make mortality estimates at advanced ages for shorter (monthly) age intervals, rather than for traditional one-year age intervals when actuarial estimate of hazard rate is applied.



**Figure 2. Deaths at extreme old ages (102 years) are not distributed uniformly over one-year age interval in contrast to earlier age groups (90 years).**

**1891 birth cohort from the SSA DMF.**

**a) Distribution for 90-year-olds**

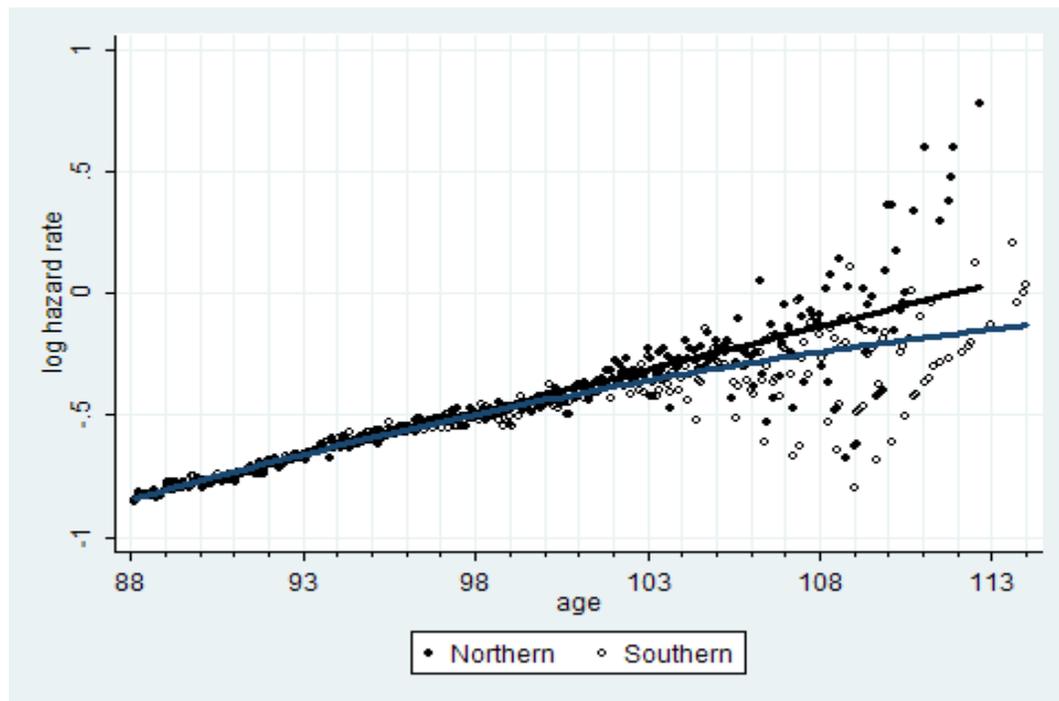
**b) Distribution for 102-year-olds**

Using single-year birth cohort data from the SSA DMF we were able to minimize the effects of population heterogeneity on age-related mortality dynamics. We were also able to calculate more accurate monthly estimates of hazard rates, which are less prone to possible biases caused by violation of typical assumptions used in mortality estimation (Kimball 1960). However we were not able to control for possible age misreporting among the deceased. The most common type of age misreporting among very old

individuals is age exaggeration (Willcox *et al.* 2008). This type of age misreporting results in underestimation of mortality rates at advanced ages. Thus, we may suggest that mortality deceleration after age 105 years among DMF cohorts may be caused by poor data quality (age exaggeration) at very advanced ages. If this hypothesis is correct then mortality deceleration at advanced ages should be less expressed for data with presumably better quality. In order to test this hypothesis we conducted a data quality study for DMF birth cohorts.

## 6. Measures to improve data quality

Mortality deceleration and even decline of mortality is observed usually for data of low quality. On the other hand, improvement of data quality results in straighter mortality trajectory in semi-log scale (Kestenbaum and Ferguson 2002). Thus, we may suggest that mortality estimates at advanced ages should be lower in populations with less accurate age reporting compared to populations with more accurate age reporting. In order to test this hypothesis, we compared mortality trajectories at advanced ages for populations with different quality of age reporting.



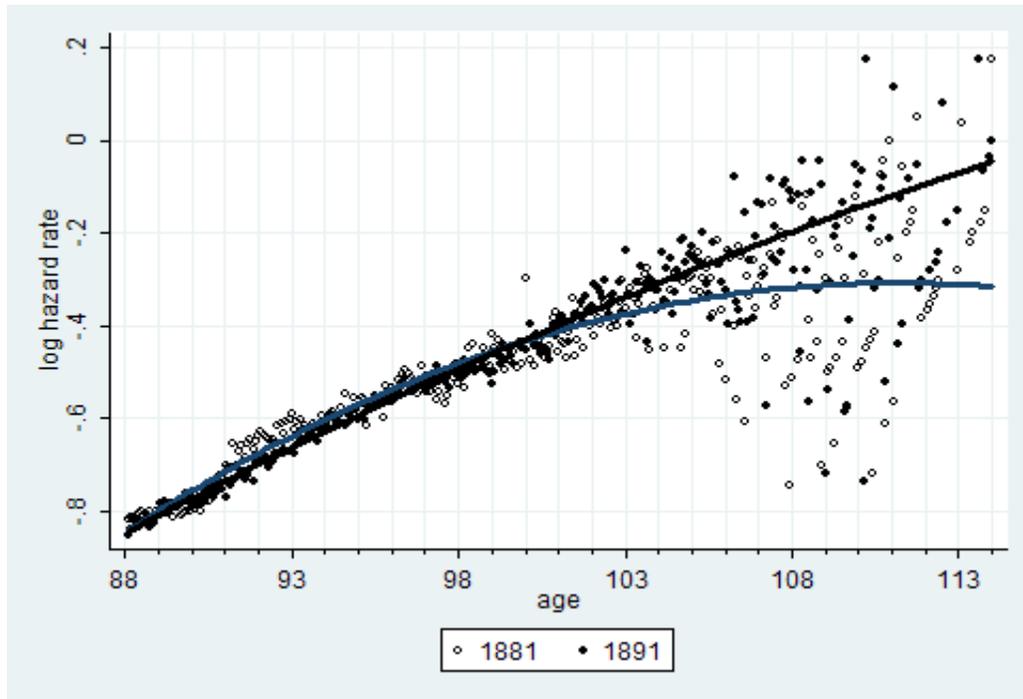
**Figure 3. Regional mortality data with presumably different quality.**

**Logarithm of hazard rate (per year) for 1891 birth cohort as a function of age. Comparison of less reliable (“Southern” group) and more reliable (“Northern” group) data. Solid lines show polynomial fit with quadratic term. Data from the Social Security Administration Death Master File.**

Study of age validation among individuals aged 110 years or more (Rosenwaike and Stone 2003) demonstrated that age reporting among supercentenarians in Social Security Administration database is rather accurate with exception of persons born in the Southern states. In order to compare populations with presumably different data quality we split DMF records into two groups. In the first (‘Southern’) group of less reliable data we included records for those persons who applied for social security number in the

Southeast states (AR, AL, GA, MS, LA, TN, FL, KY, SC, NC, VA, WV), Southwest states (AZ, NM, TX, OK), Puerto Rico and Hawaii. We also added to this group all records for persons applied to SSN in New York and California because of very high proportion of immigrants (with unknown quality of age reporting) residing in these states. In the second ('Northern') group with presumably better data quality we included persons applying to social security number in all other states as well as retired rail-road workers.

Figure 3 shows age-trajectories of mortality for persons born in 1891 and applied for SSN either in "Southern" or "Northern" states. Note that the line corresponding to polynomial fit of the logarithm of mortality with age is straighter for "Northern" group (upper line) compared to the line corresponding to mortality of "Southern" group (lower line). Thus, the estimates of age-specific mortality demonstrate straighter trajectory of mortality in semi-log scale for population of "Northern" states with presumably better quality of age reporting compared to population with less reliable quality of age reporting (verified using polynomial fit of mortality data).



**Figure 4. Historical mortality data with presumably different quality.** Logarithm of hazard rate (per year) for two birth cohorts as a function of age. Comparison of less reliable (older 1881 birth cohort, lower line) and more reliable (more recent 1891 birth cohort, upper line) data. Solid lines show fit by polynomial function with quadratic term. Data from the Social Security Administration Death Master File.

Similar phenomenon is observed when mortality data for earlier birth cohort are compared to mortality data for later birth cohort. Records for later born persons are expected to be of better quality due to improvement of age reporting over time. Taking into account that age reporting is improving over the history we may suggest that mortality for later birth cohorts would follow the Gompertz law more closely compared to the mortality of earlier birth cohorts with presumably lower data quality. This suggestion is confirmed by comparing 1881 and 1891 birth cohorts (Figure 4). More recent 1891

birth cohort demonstrates straighter trajectory and lower statistical noise after age 105 than the older 1881 one (see Figure 4). Thus, we may expect that future extinct cohorts born after 1891 may demonstrate even better fit by the Gompertz model than the older ones because of continuously improving accuracy in age reporting.

Quantitative estimation of the degree of mortality deceleration at older ages was conducted by modeling the logarithm of hazard rate as a polynomial function of age. In the case of mortality deceleration, the hazard rate would be described by a convex parabola in semi-log coordinates. This is expressed as a negative parameter at quadratic term and the absolute value of this parameter can be used as a measure of mortality deceleration. Table 1 presents values of the parameter at quadratic term for polynomial regression of logarithm of hazard rate as a function of age (after age 90). Note that the absolute value of parameter at quadratic term is higher for older birth cohorts and 'Southern' group of less reliable records (described earlier) confirming that data with lower quality demonstrate a higher degree of mortality deceleration.

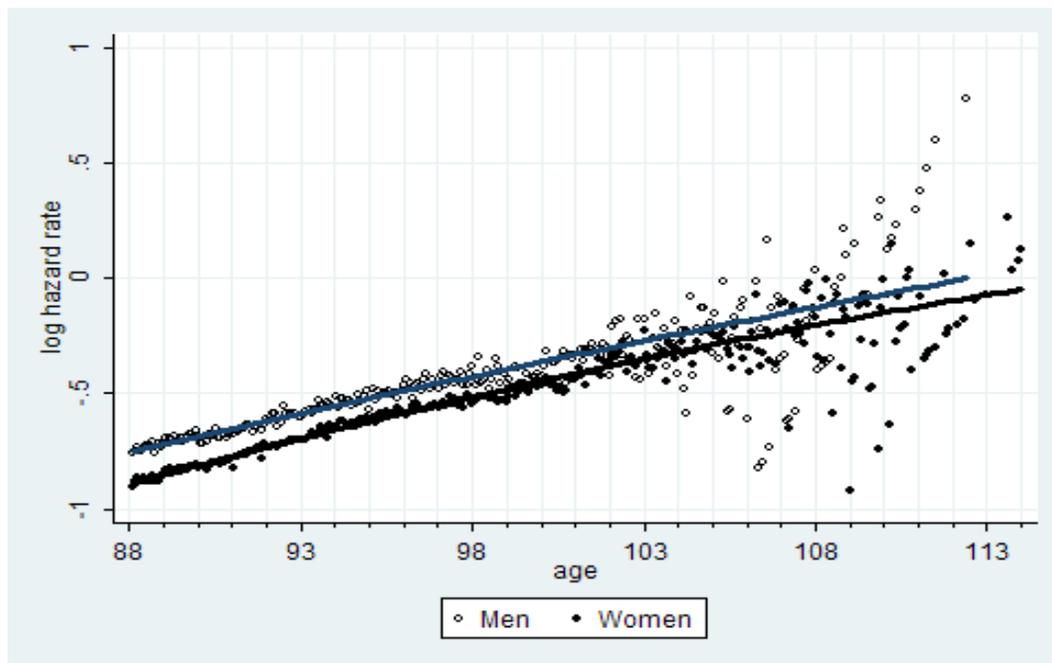
**Table 1. Comparison of the degree of mortality deceleration at advanced ages using polynomial regression model<sup>†</sup>.**

Population	Parameter at quadratic term x 10 <sup>5</sup> (95%CI)
1881 birth cohort	-75.5 (-76.7, -74.4)
1889 birth cohort	-38.3 (-38.9, -37.8)
1894 birth cohort	5.2 (4.7, 5.6)
1891 birth cohort:	
Northern group (more reliable data)	-18.7 (-19.6, -17.8)
Southern group (less reliable data)	-64.7 (-65.7, -63.6)
1893 birth cohort:	
Northern group (more reliable data)	-7.4 (-8.2, -6.6)
Southern group (less reliable data)	-18.3 (-19.2, -17.3)
1891 birth cohort:	
yearly age intervals	-48.9 (-49.1, -48.6)
monthly age intervals	-25.5 (-25.8, -25.1)

<sup>†</sup> Logarithm of hazard rate after age 90 for particular population was fitted by polynomial regression model with quadratic term.

Based on the results presented above we may conclude that the better the quality of mortality data at advanced ages the closer mortality trajectories are to the Gompertz function. However, we have to admit that the data at very old ages are noisy and have quality problems even for 'Northern' group and the most recent birth cohorts. Thus, our next step in data quality control is to identify the age interval with reasonably good data quality and then to compare competing mortality models in this age interval.

One approach to test data quality at advanced ages is to calculate female to male ratio at advanced ages. Taking into account that female mortality is always lower than male mortality it is reasonable to expect that the female-to-male ratio should continuously increase with age. On the other hand, old men have a tendency for age exaggeration and in populations with poor age registration there is a relative excess of men at very advanced ages (Caselli *et al.* 2006; Willcox *et al.* 2008).

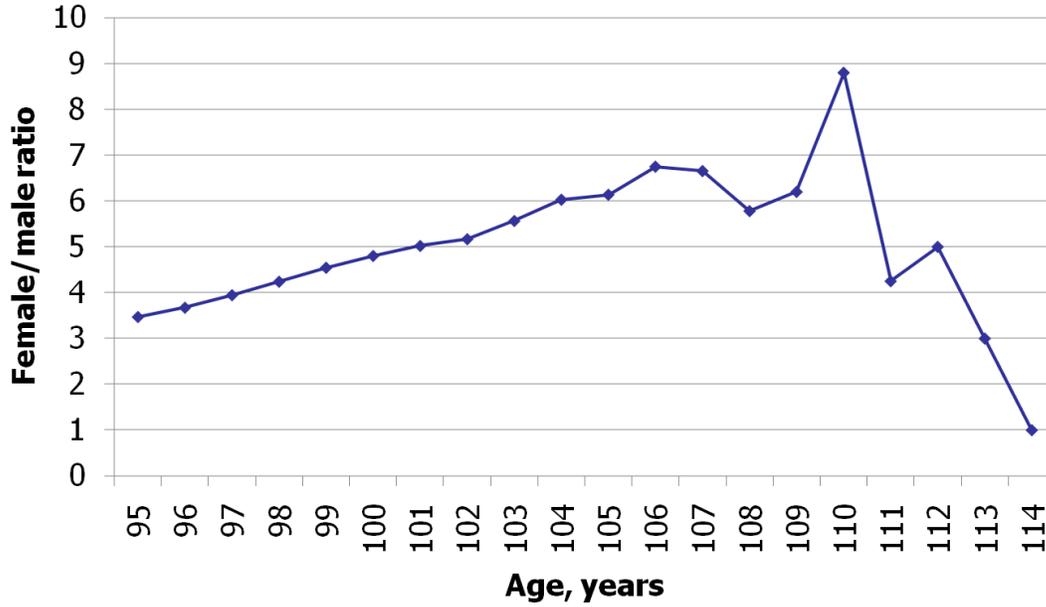


**Figure 5. Age-specific hazard rates (log scale) for men (upper line) and women (lower line) born in 1891. Solid lines show fit with polynomial function with quadratic term. Data from the Social Security Administration Death Master File.**

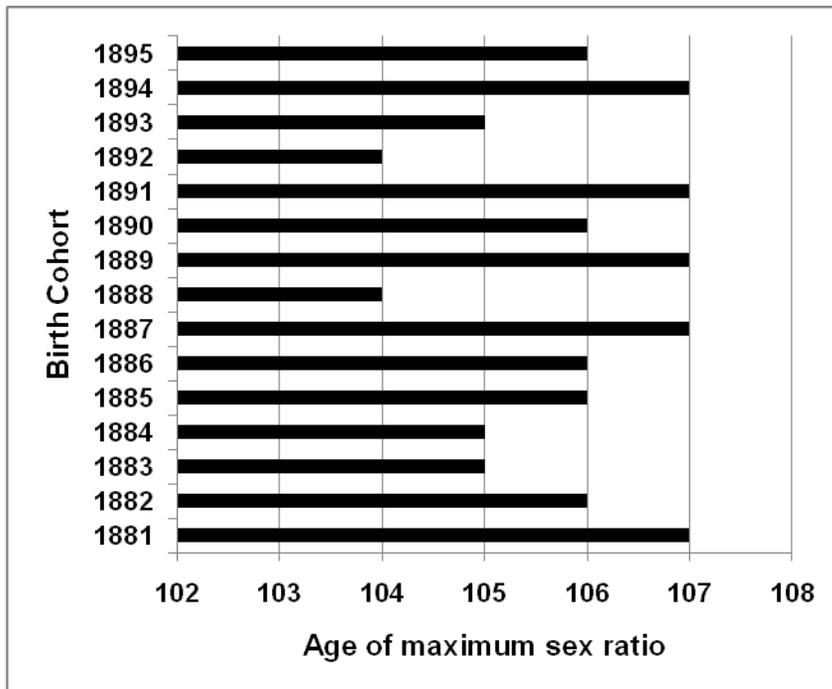
The SSA DMF does not have information about sex of the deceased. To cope with this limitation of data sample, we conducted a procedure of sex identification using information about the 1,000 most commonly used baby first names in the 1900s provided by the Social Security Administration (<http://www.ssa.gov/OACT/babynames>). These data come from a sample of 5 percent of all Social Security cards issued to individuals who were born during the 1900s in the United States. From the lists of male and female names we removed names consisting of initials and potentially unisex names (like Jessie or Lonnie). It is interesting to note that the SSA male list contains some obviously female names (Mary, Elizabeth) and the same problem was observed for the female list, which indicates that the SSA data apparently contain many sex misidentifications. These female names were removed from the male list and the same procedure was done for the female list. Additional male or female names found in DMF were also added to our sex identification name lists. Using the final lists of male and female first names we applied the sex identification procedure to the sample of 1879-1895 birth cohorts. Eventually we were able to identify sex for 91-93 percent of records for persons in our sample. The remaining 7-9 percent of persons with unknown sex had approximately the same mean lifespan as the remaining percent of individuals with identified sex pooled together (checked with the t-test statistics), so the existence of possible sex non-identification bias in mortality seems unlikely. This data sample of 774,926 individuals with known sex (out of 846,982 individuals) was used for identifying the age interval of reasonably good data quality using information on the female/male ratio at advanced ages.

The result of hazard rate estimation for men and women born in 1891 is presented in Figure 5. Note that male mortality continues to exceed female mortality up to very advanced ages and this mortality differential narrows very slowly with age. At age 110 years the number of remaining males (9) and females (61) is too small for accurate

estimates of hazard rate after this age. This pattern of sex differential in mortality at very old ages confirms the validity of using female-to-male ratio as a tool of data quality assessment because higher male mortality at all observed ages should theoretically lead to continuously growing female/male ratio.



**Figure 6.** Observed female to male ratio at advanced ages for combined 1887-1892 birth cohort. If data are of good quality then this ratio should grow with age



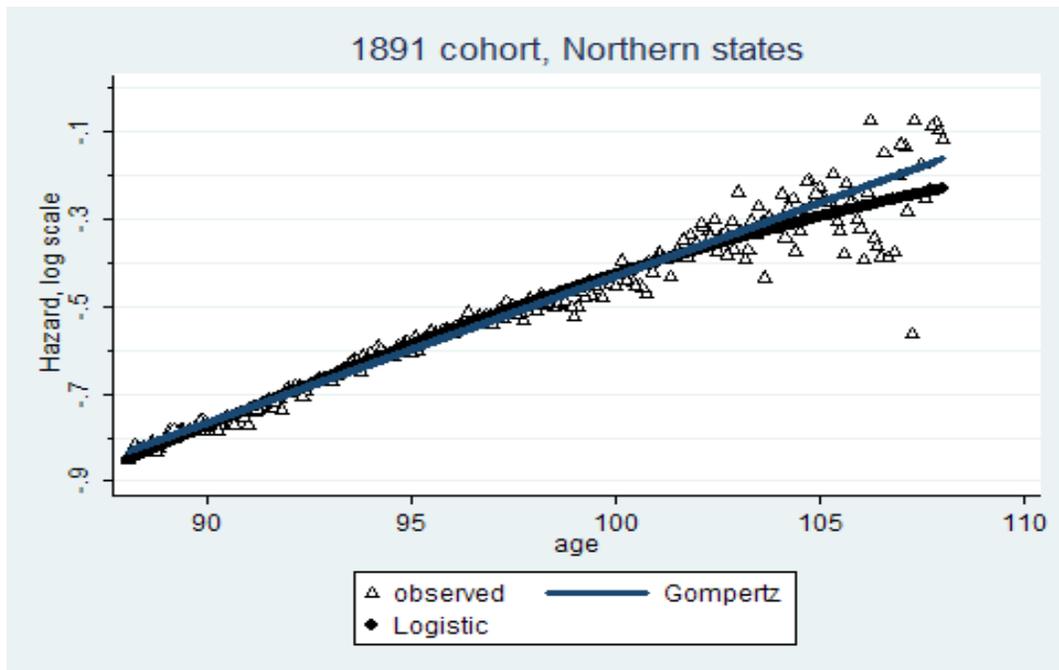
**Figure 7.** Age of maximum female to male ratio, by birth cohort

We calculated female-to-male ratio after age 95 years for 1881-1895 U.S. birth cohorts from the SSA DMF. Figure 6 demonstrates the age dependency of this ratio for pooled sample of 1887-1892 birth cohorts (these cohorts have similar levels of mortality). Note that the female to male ratio is growing steadily with age up to ages 106-107 years. After this age the female-to-male ratio starts to decrease indicating declining quality of age reporting. Thus, the estimates of hazard rates obtained from the SSA DMF are of acceptable quality up to the age of 106 years. Figure 7 shows the maximum age or tipping point when the female-to-male ratio starts to decline for different birth cohorts. It demonstrates that this age varies from 104 to 107 indicating the upper limit for data of reasonably good quality in terms of age reporting.

Thus, our study of female-to-male ratio for different birth cohorts indicates that age reporting has reasonably good quality up to the age of about 106 years. For this reason we used age interval 88-106 years for mortality modeling.

### 7. Modeling mortality at advanced ages

Next step of our study was to compare two competing models of mortality at advanced ages - the Gompertz and the logistic models - using data of reasonably good quality. Study of data quality control at advanced ages described earlier suggests that age reporting among the oldest-old in the United States is good until the age of 106 years. It means that comparing mortality models beyond this age is not feasible because of poor quality of mortality data. For this reason, we used a subsample of deaths for persons applied to SSNs in the 'Northern' states (described above) and born in 1886-1895, because these data have reasonably good quality. We applied Gompertz and logistic (Kannisto) models (Thatcher et al. 1998) to mortality modeling in the age interval 88-106 years using nonlinear regression method for parameter estimation. Calculations were performed using Stata statistical software, release 11.



**Figure 8. Mortality fitting using Gompertz (upper line) and logistic (lower line) models for 1891 birth cohort of persons applied to SSNs in the Northern states. Data from the Social Security Administration Death Master File**

Figure 8 presents the results of mortality fitting after age 88 for Gompertz and Kannisto models using data for 'Northern' states and 1891 birth cohort. Visual impression is that Gompertz model fits data better than the logistic model although for objective comparisons we need to use conventional measures of goodness-of-fit. In this study, goodness-of-fit for each model was estimated using the Bayesian information criterion (BIC). Table 2 shows values of BIC for both Gompertz and logistic model for ten studied birth cohorts. Note that in 8 out of 10 cases (studied birth cohorts), the Gompertz model demonstrates better fit (lower BIC) than the logistic model for age interval 88-106 years.

**Table 2. Comparison of goodness-of-fit (Bayesian Information Criterion, BIC) for Gompertz and logistic models of mortality <sup>†</sup>**

Birth cohort	Cohort size at age 88 years, persons	Bayesian Information Criterion (BIC)	
		Gompertz model <sup>††</sup>	Logistic model
1886	111,657	<b>-594776.2</b>	-588049.5
1887	114,469	<b>-625303.0</b>	-618721.4
1888	128,768	-709620.7	-712575.5
1889	131,778	-710871.1	-715356.6
1890	135,393	<b>-724731.0</b>	-722939.6
1891	143,138	<b>-767138.3</b>	-739727.6
1892	152,058	<b>-831555.3</b>	-810951.8
1893	156,189	<b>-890022.6</b>	-862135.9
1894	160,835	<b>-946219.0</b>	-905787.1
1895	165,294	<b>-921650.3</b>	-863246.6

<sup>†</sup> Estimates were made in the age interval 88-106 years for ten single-year U.S. birth cohorts and data of enhanced accuracy for individuals applied to SSNs in the Northern states (see explanation in the text).

<sup>††</sup> Cases when the Gompertz model fits data better than the logistic model are highlighted in bold.

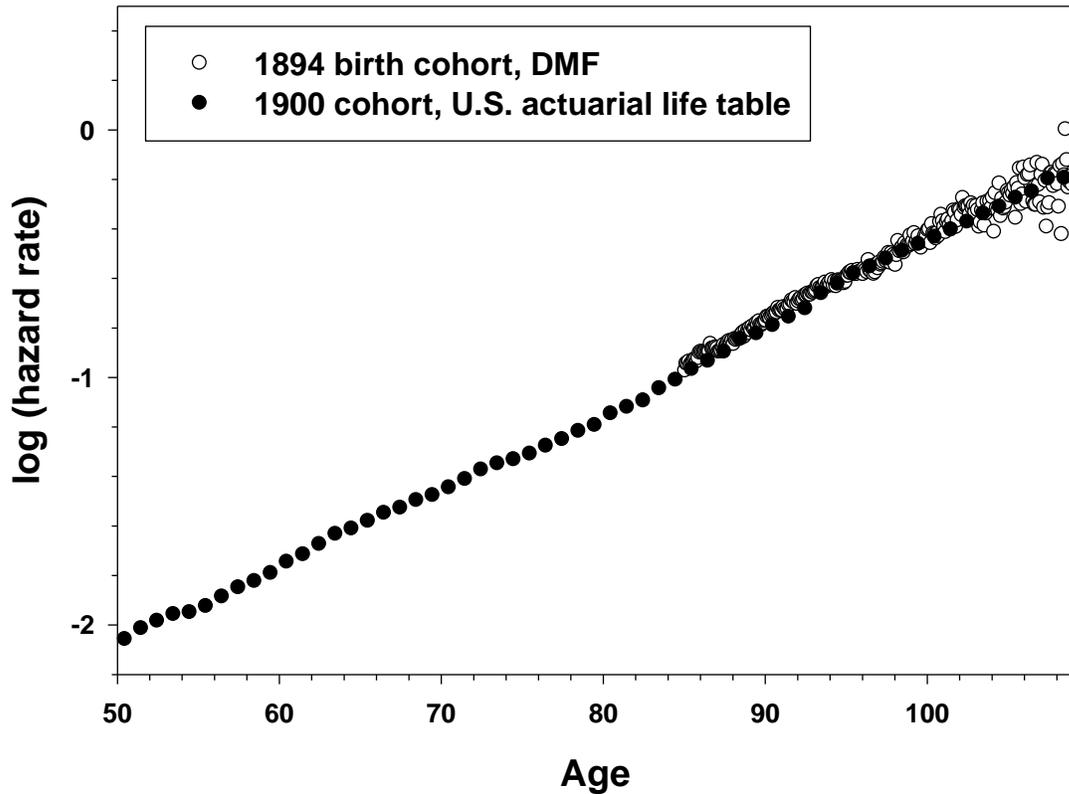
At this moment, we cannot make a conclusion that Gompertz model fits mortality data better than the logistic model after age 106, because of low quality of age reporting beyond this age. At the same time, the data indicate that the Gompertz model fits mortality data well until the age 106 years. Taking into account that survival beyond age 106 years is rather rare event, it would be reasonable to suggest the use of Gompertz model rather than logistic model for closing cohort life tables in demographic practice. In this case, mortality modeling could be done first for hazard rate (mortality force) function and then all life table functions (including probability of death,  $q_x$ ) could be derived on the basis of modeled values of hazard rate.

## 8. Comparison of DMF old-age mortality data with Actuarial Life Table

Due to some limitations of the DMF data (incompleteness of death registration before the 1970s), we were unable to estimate mortality before the age 85-88 years. There is also a question whether our estimates of Gompertz parameters (slope parameter in particular) are applicable to a wider age interval or they are specific only for advanced ages (so-called two-stage Gompertz model). In this study we compared our empirical estimates of cohort mortality at advanced ages, based on DMF data, with 1900 actuarial cohort life table (Bell, Wade and Goss 1992). Unfortunately, there are no reliable U.S. cohort life tables for cohorts older than 1900, so we compared our mortality data for earlier birth cohorts to later 1900 birth cohort. It should be noted that adult U.S. population did not experience substantial mortality changes between 1890 and 1900.

According to the estimates made by historical demographers, life expectancy at age 40 in the United States in 1890 (27.61 years) was practically the same as in 1900 (27.52 years) (Haines 1998). Thus, we believe that mortality at advanced ages for 1900 birth cohort in the actuarial study and 1894 birth cohort we used for comparison should not be substantially different.

We initially compared our DMF data with probability of death estimates from actuarial life table ( $q_x$ ) and found that life table values of probability of death show deceleration after age 100. It should be noted, however, that the yearly estimates of probability of death ( $q_x$ ) and mortality force ( $\mu_x$ ) diverge at advanced ages. This happens because probability of death cannot continue its rapid growth when it approaches its mathematical upper limit equal to 1, while mortality force theoretically can increase indefinitely with age. Therefore, estimates of probability of death at advanced ages are biased downward compared to the hazard rate estimates. Hence, probability of death estimates beyond age 100 show deceleration in semi-log coordinates even if values of hazard rates follow the Gompertz model.



**Figure 9 . Actuarial 1900 U.S. cohort life table and SSDI 1887 birth cohort. Source for actuarial life table: Bell, F.C., Miller, M.L. Life Tables for the United States Social Security Area 1900-2100. Actuarial Study No. 116.**

Formula proposed by George Sacher (Sacher 1956; Sacher 1966) gives more accurate estimates of hazard rate for yearly age intervals, so we applied a simplified Sacher formula for age-specific hazard rate conversion from  $q_x$  values reported in 1900 actuarial cohort life table:

$$\mu_x = -\ln(1-q_x)$$

where  $q_x$  is a probability of death value from the life table.

Figure 9 shows the trajectories of age-specific hazard rates for 1900 birth cohort from the actuarial study and 1894 birth cohort from DMF over a broad age interval starting at age 50 years. Note that hazard rate estimates for the 1894 birth cohort are practically identical to the hazard rate estimates calculated on the basis of the actuarial life table. Also note that mortality of 1894 birth cohort has the same slope in semi-log coordinates as mortality of actuarial birth cohort calculated for much wider age interval, which does not support the suggestion about the two-stage Gompertz model of mortality at advanced ages. Indeed, the maximum likelihood estimator of the Gompertz slope parameter for mortality in 1894 cohort measured in the interval 88-106 years (0.0786 year<sup>-1</sup>, 95%CI: 0.0786-0.0787) does not differ from the slope parameter calculated over the age interval 40-104 years in 1900 life table: 0.0785 year<sup>-1</sup>, 95%CI: 0.0772-0.0797. Thus, we may conclude that more accurate estimates of hazard rates at advanced ages based on individual mortality data practically coincide with hazard rate estimates calculated on the basis of 1900 actuarial cohort life table (Bell *et al.* 1992). It is remarkable that estimates in actuarial life table obtained by extrapolation of annual probabilities of death after age 95 (Bell *et al.* 1992) are practically identical to the observed estimates of hazard rates based on mortality experience of very large cohort of old individuals.

## 9. Discussion

This study of large single-year U.S. birth cohorts found that mortality deceleration at advanced ages is negligible up to the age of 106 years. Below the age of 107 years and for data of reasonably good quality the Gompertz model fits mortality better than the logistic model (no mortality deceleration). We also found that deceleration of mortality in later life is more expressed for data of lower quality. Quality of age reporting in DMF becomes particularly poor after the age of 107 years. It is also interesting that the DMF mortality data agree remarkably well with mortality estimates reported by the 1900 actuarial cohort life table (if hazard rate values are compared).

There are several reasons why earlier studies, including our own research (Gavrilov 1984; Gavrilov and Gavrilova 1991), reported mortality deceleration and mortality leveling-off at advanced ages (Horiuchi and Wilmoth 1998; Kannisto 1994; Robine and Vaupel 2001; Thatcher 1999; Thatcher *et al.* 1998). First, many studies present information for age-specific probability of death rather than hazard rate (Gallop and Macdonald 2005; Robine and Vaupel 2001). It is not surprising that probability of death has a tendency of deceleration at advanced ages when mortality is high, taking into account that this mortality indicator has theoretical upper limit equal to one. For example, a study of mortality among supercentenarians demonstrated that probability of death for this group does not increase with age (Robine and Vaupel 2001). The authors do not provide estimates of hazard rate for this small heterogeneous population, so at this moment it is difficult to make conclusion about the real mortality trajectory in this sample of very old individuals. In the study of validated French-Canadian centenarians born in 1870-1894 the authors found no growth of mortality with age for life table death rates rather than hazard rates (Beaudry-Godin, Bourbeau and Desjardins 2008). It is true that for young and middle ages (when mortality is relatively low) probability of death and hazard rate values are numerically close to each other. As a result, some authors do not

distinguish between probability of death and hazard rate in their calculations (Le Bras 2005).

In the studies of period mortality, deceleration at advanced ages is probably caused by mixing together data for several birth cohorts with different mortality levels (heterogeneity effect) and using cross-sectional instead of cohort data (Horiuchi and Wilmoth 1998; Kannisto 1994; Thatcher et al. 1998). In addition to that, if information about population at risk is taken from censuses there is always a possibility of mismatch in the accuracy of age reporting between deaths and population at risk due to higher proportion of age misreporting in censuses (so-called denominator problem). It should be noted that mortality deceleration was also reported for cohort mortality data (Thatcher et al. 1998). In this case, mortality deceleration may be caused by age misreporting in death data for older persons as we already found in our study. Earlier studies, conducted more than ten years ago, used data for older birth cohorts when age reporting was not particularly accurate even for such countries as the U.K (Gallop and Macdonald 2005). In addition to that, most developed countries have much smaller populations compared to the United States and hence studies of mortality at advanced ages for these countries have to combine together many single-year birth cohorts thereby increasing the heterogeneity of the sample. Thus, age exaggeration, use of probabilities of death instead of hazard rates and perhaps data heterogeneity could lead to downward biases in mortality estimates at older ages reported in previous studies.

In our study we found no significant mortality deceleration at advanced age for humans while many lower organisms demonstrate very long late-age mortality plateaus (Carey et al. 1992; Curtsinger et al. 1992; Gavrilov and Gavrilova 1991; Vaupel et al. 1998). Earlier studies showed that the period of mortality deceleration in mammalian species is very short (Lindop 1961; Sacher 1966) compared to lower organisms (Gavrilov and Gavrilova 2006; Vaupel et al. 1998). It appears to be relatively short if not negligible in humans too. This observation agrees with the prediction of the reliability theory of aging that more complex living systems/organisms with many vital subsystems (like mammals) may experience very short period of mortality plateau at advanced ages in contrast to more simple living organisms (Gavrilov and Gavrilova 1991; Gavrilov and Gavrilova 2001; Gavrilov and Gavrilova 2006). Thus, human mortality at advanced ages may show very short or even negligible mortality plateau despite the evidence of mortality leveling-off among lower organisms such as insects.

The results obtained in this study may be important for demographic practice, particularly if mortality is analyzed for real birth cohorts. For cohort data, hazard rates can be extrapolated with the Gompertz formula up to 107 years of age and then probabilities of death can be reconstructed with Sacher formula if necessary. Our study also suggests that improvement of age reporting in history may result in mortality trajectories that better follow the Gompertz function until very advanced ages.

### **Acknowledgements**

This study was supported by the National Institute on Aging (NIA grant R01 AG028620).

## References

- Austad, S.N. 2001. "Concepts and theories of aging." Pp. 3-22 in *Handbook of the biology of aging*, edited by E.J. Masoro and S.N. Austad. San Diego: Academic Press.
- Beard, R.E. 1959. "Note on some mathematical mortality models." Pp. 302-311 in *The lifespan of animals*, edited by E.W. Wolstenholme and M.O. O'Connor. Boston: Little, Brown.
- . 1971. "Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes." Pp. 57-68 in *Biological Aspects of Demography*, edited by W. Brass. London: Taylor and Francis.
- Beaudry-Godin, M., R. Bourbeau, and B. Desjardins. 2008. "Data validation and measurement of cohort mortality among centenarians in Quebec (Canada) according to ethnic origin." in *Living to 100 and Beyond [online monograph]*. Shaumburg: Society of Actuaries.
- Bell, F.C., A.H. Wade, and S.C. Goss. 1992. *Life Tables for the United States Social Security Area 1900-2080. Actuarial Study No.107*. Baltimore, MD: U. S. Department of Health and Human Services.
- Bronikowski, A.M., S.C. Alberts, J. Altmann, C. Packer, K.D. Carey, and M. Tatar. 2002. "The aging baboon: comparative demography in a non-human primate." *Proceedings of the National Academy of Sciences of the USA* 99:9591-9595.
- Carey, J.R., P. Liedo, D. Orozco, and J.W. Vaupel. 1992. "Slowing of mortality-rates at older ages in large medfly cohorts." *Science* 258(5081):457-461.
- Caselli, G., L. Pozzi, J.W. Vaupel, L. Deiana, G. Pes, C. Carru, C. Franceschi, and G. Baggio. 2006. "Family clustering in Sardinian longevity: A genealogical approach." *Experimental Gerontology* 41(8):727-736.
- Charlesworth, B. 2001. "Patterns of age-specific means and genetic variances of mortality rates predicted by mutation-accumulation theory of ageing." *Journal of Theoretical Biology* 210:47-65.
- Cleves, M.A., R.G. Gutierrez, W.W. Gould, and Y. Marchenko. 2008. *An Introduction to Survival Analysis Using Stata*. College Station, Texas: Stata Press.
- Curtsinger, J.W., H. Fukui, D. Townsend, and J.W. Vaupel. 1992. "Demography of genotypes: Failure of the limited life-span paradigm in *Drosophila melanogaster*." *Science* 258:461-463.
- Curtsinger, J.W., N.S. Gavrilova, and L.A. Gavrilov. 2006. "Biodemography of Aging and Age-Specific Mortality in *Drosophila melanogaster*." Pp. 261-288 in *Handbook of the Biology of Aging*, edited by E.J. Masoro and S.N. Austad. San Diego: Academic Press.
- Depoid, F. 1973. "Mortality of old people over 85." *Population* 28(4-5):755-792.
- Economos, A.C. 1979. "A non-gompertzian paradigm for mortality kinetics of metazoan animals and failure kinetics of manufactured products." *AGE* 2:74-76.
- . 1980. "Kinetics of metazoan mortality." *Journal of Social and Biological Structures* 3(4):317-329.
- . 1983. "Rate of aging, rate of dying and the mechanism of mortality." *Archives in Gerontology and Geriatrics* 1:3-27.
- . 1985. "Rate of aging, rate of dying and non-Gompertzian mortality - encore." *Gerontology* 31(2):106-111.
- Faig, K. 2001. "Reported deaths of centenarians and near-centenarians in the U.S. Social Security Administration's Death Master File." Presented at The Society of Actuaries "Living to 100 and Beyond International Symposium", Orlando, FL.

- Gallop, A. and A.S. Macdonald. 2005. "Mortality at advanced ages in the United Kingdom." in *Living to 100 and Beyond Monograph [online edition]*. . Shaumburg: Society of Actuaries.
- Gavrilov, L.A. 1980. *Study of Life Span Genetics Using the Kinetic Analysis. PhD Thesis*. Moscow: Moscow State University, Russia.
- . 1984. "Does the limit of the life-span really exist?" *Biofizika* 29(5):908-909.
- Gavrilov, L.A. and N.S. Gavrilova. 1991. *The Biology of Life Span: A Quantitative Approach*. New York: Harwood Academic Publisher.
- . 2001. "The reliability theory of aging and longevity." *Journal of Theoretical Biology* 213(4):527-545.
- . 2006. "Reliability Theory of Aging and Longevity." Pp. 3-42 in *Handbook of the Biology of Aging*, edited by E.J. Masoro and S.N. Austad. San Diego: Academic Press.
- Gavrilova, N.S. and L.A. Gavrilov. 2007. "Search for Predictors of Exceptional Human Longevity: Using Computerized Genealogies and Internet Resources for Human Longevity Studies." *North American Actuarial Journal* 11(1):49-67.
- Gehan, E.A. 1969. "Estimating survival functions from life table." *Journal of Chronic Diseases* 21(9-10):629-8.
- Gehan, E.A. and M.M. Siddiqui. 1973. "Simple regression methods for survival time studies." *Journal of the American Statistical Association* 68(344):848-856.
- Gompertz, B. 1825. "On the nature of the function expressive of the law of human mortality and on a new mode of determining life contingencies." *Philos. Trans. Roy. Soc. London A* 115:513-585.
- Greenwood, M. and J.O. Irwin. 1939. "The biostatistics of senility." *Human Biology* 11:1-23.
- Haines, M.R. 1998. "Estimated life tables for the United States, 1850-1910." *Historical Methods* 31(4):149-169.
- Horiuchi, S. and J.R. Wilmoth. 1998. "Deceleration in the age pattern of mortality at older ages." *Demography* 35:391-412.
- Kannisto, V. 1988. "On the survival of centenarians and the span of life." *Population Studies-a Journal of Demography* 42:389-406.
- . 1994. *Development of Oldest-Old Mortality, 1950-1990: Evidence from 28 Developed Countries*. Odense: Odense University Press.
- Kestenbaum, B. and B.R. Ferguson. 2002. "Mortality of the extreme aged in the United States in the 1990s, based on improved Medicare data." in *Living to 100 and Beyond International Symposium*. Shaumburg: The Society of Actuaries.
- Kimball, A.W. 1960. "Estimation of mortality intensities in animal experiments." *Biometrics* 16(4):505-521.
- Klein, J.P. and M.L. Moesberger. 1997. *Survival analysis techniques for censored and truncated data*. New York: Springer-Verlag.
- Le Bras, H. 1976. "Lois de mortalité et age limité." *Population* 31:655-692.
- . 2005. "Mortality tempo versus removal of causes of mortality: opposite views leading to different estimations of life expectancy." *Demographic Research* 13:615-640.
- Lindop, P.J. 1961. "Growth rate, lifespan and causes of death in SAS/4 mice." *Gerontologia* 5:193-208.
- Mueller, L.D. and M.R. Rose. 1996. "Evolutionary theory predicts late-life mortality plateaus." *Proceedings of the National Academy of Sciences, USA* 93:15249-15253.
- Olshansky, S.J. 1998. "On the biodemography of aging: A review essay." *Population and Development Review* 24:381-393.
- Robine, J.-M., A. Cournil, J. Gampe, and J.W. Vaupel. 2005. "IDL, the International Database on Longevity." Society of Actuaries.

- Robine, J.M. and J.W. Vaupel. 2001. "Supercentenarians: slower ageing individuals or senile elderly?" *Experimental Gerontology* 36(4-6):915-930.
- Rosenwaikie, I. and L.F. Stone. 2003. "Verification of the ages of supercentenarians in the United States: Results of a matching study." *Demography* 40(4):727-739.
- Sacher, G.A. 1956. "On the statistical nature of mortality, with especial reference to chronic radiation mortality." *Radiology* 67:250-257.
- . 1966. "The Gompertz transformation in the study of the injury-mortality relationship: Application to late radiation effects and ageing." Pp. 411-441 in *Radiation and Aging*, edited by P.J. Lindop and G.A. Sacher. London: Taylor and Francis.
- Sesso, H.D., R.S. Paffenbarger, and I.M. Lee. 2000. "Comparison of National Death Index and World Wide Web death searches." *American Journal of Epidemiology* 152(2):107-111.
- StataCorp. 2009. *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.
- Strehler, B.L. and A.S. Mildvan. 1960. "General theory of mortality and aging." *Science* 132(3418):14-21.
- Tatar, M., J.R. Carey, and J.W. Vaupel. 1993. "Long-term cost of reproduction with and without accelerated senescence in *Callosobruchus maculatus*: Analysis of age-specific mortality." *Evolution* 47:1302-1312.
- Thatcher, A.R. 1999. "The long-term pattern of adult mortality and the highest attained age." *Journal of the Royal Statistical Society Series a-Statistics in Society* 162:5-30.
- Thatcher, A.R., V. Kannisto, and J.W. Vaupel. 1998. *The Force of Mortality at Ages 80 to 120*. Odense: Odense University Press.
- Vaupel, J.W., J.R. Carey, K. Christensen, T.E. Johnson, A.I. Yashin, N.V. Holm, I.A. Iachine, V. Kannisto, A.A. Khazaeli, P. Liedo, V.D. Longo, Y. Zeng, K.G. Manton, and J.W. Curtsinger. 1998. "Biodemographic trajectories of longevity." *Science* 280(5365):855-860.
- Vaupel, J.W., K.G. Manton, and E. Stallard. 1979. "Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." *Demography* 16(3):439-454.
- Vincent, P. 1951. "La mortalite des Viellards." *Population* 6:181-204.
- Watson, G.S. and M.R. Leadbetter. 1964. "Hazard analysis .I." *Biometrika* 51(1-2):175-&.
- Willcox, D.C., B.J. Willcox, Q. He, N.C. Wang, and M. Suzuki. 2008. "They really are that old: A validation study of centenarian prevalence in Okinawa." *Journals of Gerontology Series a-Biological Sciences and Medical Sciences* 63(4):338-349.