The growing number of persons living beyond age 80 underscores the need for accurate measurement of mortality at advanced ages.

Recent projections of the U.S. Census Bureau significantly overestimated the actual number of centenarians.

New estimates based on the 2010 census are two times lower than the U.S. Bureau of Census forecast.

The same story recently happened in the Great Britain.
Mortality at advanced ages is the key variable for understanding population trends among the oldest-old.

The first comprehensive study of mortality at advanced ages was published in 1939.

Earlier studies suggested that the exponential growth of mortality with age (Gompertz law) is followed by a period of deceleration, with slower rates of mortality increase.

M. Greenwood, J. O. Irwin. BIOSTATISTICS OF SENILITY

...possibility that with advancing age the rate of mortality asymptotes to a finite value. *

...The limiting values of $\delta_0$ are 0.439 for women and 0.544 for men. Some tests of the ultimate mortalities in non-human experience were not unfavorable. *
Mortality at Advanced Ages, Recent Study

![Graph](image1)


Existing Explanations of Mortality Deceleration

- **Population Heterogeneity** (Beard, 1959; Sacher, 1966). "... sub-populations with the higher injury levels die out more rapidly, resulting in progressive selection for vigour in the surviving populations" (Sacher, 1966)
- **Exhaustion of organism’s redundancy** (reserves) at extremely old ages so that every random hit results in death (Gavrilov, Gavrilova, 1991; 2001)
- **Lower risks of death for older people** due to less risky behavior (Greenwood, Irwin, 1939)
- **Evolutionary explanations** (Mueller, Rose, 1996; Charlesworth, 2001)

Mortality force (hazard rate) is the best indicator to study mortality at advanced ages

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

- Does not depend on the length of age interval
- Has no upper boundary and theoretically can grow unlimitedly
- Famous Gompertz law was proposed for fitting age-specific mortality force function (Gompertz, 1825)

Problems in Hazard Rate Estimation At Extremely Old Ages

1. Mortality deceleration in humans may be an artifact of 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
3. Ages of very old people may be highly exaggerated

Monthly Estimates of Mortality are More Accurate

Simulation assuming Gompertz law for hazard rate

State package uses the Nelson-Aalen estimate of hazard rate:

\[ \mu_x = H(x) - H(x-1) \cdot \frac{d}{dx} \]

\[ H(x) \text{ is a cumulative hazard function, } d \text{ is the number of deaths occurring at time } x \text{ and } n_x \text{ is the number at risk at time } x \text{ before the occurrence of the deaths. This method is equivalent to calculation of probabilities of death:} \]

\[ q_x = \frac{d}{n_x} \]

Social Security Administration’s Death Master File (SSA’s DMF) Helps to Alleviate the First Two Problems

- Allows to study mortality in large, more homogeneous single-year or even single-month birth cohorts
- Allows to estimate mortality in one-month age intervals narrowing the interval of hazard rates estimation
What Is SSA’s DMF?

- As a result of a court case under the Freedom of Information Act, SSA is required to release its death information to the public. SSA’s DMF contains the complete and official SSA database extract, as well as updates to the full file of persons reported to SSA as being deceased.
- SSA DMF is no longer a publicly available data resource (now is available from Ancestry.com for fee)
- We used DMF full file obtained from the National Technical Information Service (NTIS). Last deaths occurred in September 2011.

SSA’s DMF Advantage

- Some birth cohorts covered by DMF could be studied by the method of extinct generations
- Considered superior in data quality compared to vital statistics records by some researchers

Social Security Administration’s Death Master File (DMF) Was Used in This Study:

To estimate hazard rates for relatively homogeneous single-year extinct birth cohorts (1890-1899)
To obtain monthly rather than traditional annual estimates of hazard rates
To identify the age interval and cohort with reasonably good data quality and compare mortality models

Hypothesis

Mortality deceleration at advanced ages 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 better quality
Quality Control (1)

Study of mortality in the states with different quality of age reporting:
Records for persons applied to SSN in the Southern states were found to be of lower quality (Rosenwaike, Stone, 2003)
We compared mortality of persons applied to SSN in Southern states, Hawaii, Puerto Rico, CA and NY with mortality of persons applied in the Northern states (the remainder)

Quality Control (2)

Study of mortality for earlier and later single-year extinct birth cohorts:
Records for later born persons are supposed to be of better quality due to improvement of age reporting over time.

New Pilot Study based on DMF:
Mortality of Railroad Retirees (SSN: 700-728)

- In the past railroad workers could have better age reporting compared to their peers
- If mortality deceleration is caused by age misreporting, then better data quality for railroad workers may lead to less mortality deceleration among them
Mortality of Railroad Retirees and their non-Railroad Peers
Males, 1895-99 birth cohort

Straight lines correspond to the quadratic fit of hazard rates in semi-log coordinates
For RR group, coefficient at quadratic term is positive and significant; for non-RR group this coefficient is not significant

At what age interval data have reasonably good quality?

A study of age-specific mortality by gender

Women have lower mortality at advanced ages

Hence number of females to number of males ratio should grow with age

Rapid deterioration of data quality with age in SSA DMF

Research Article
Typologies of Extreme Longevity Myths

Robert De Young,1 Berton P. Beunen2,3,4 Kévin McLaughlin5,6,7, Michel Poutouia7 and Thomas E. Perls

- "Invalid age claim rates increase with age from 65% at age 110-111 to 98% by age 115 to 100% for 120+ years."
  (Current Gerontology and Geriatrics Research, Vol.2010, Article ID 423087)
Selection of competing mortality models using DMF data

- Data with reasonably good quality were used: non-Southern states and 85-106 years age interval
- Gompertz and logistic (Kannisto) models were compared
- Nonlinear regression model for parameter estimates (Stata 11)
- Model goodness-of-fit was estimated using AIC and BIC

Akaike information criterion (AIC) to compare Kannisto and Gompertz models, **men**, by birth cohort (non-Southern states)

Conclusion: In all ten cases Gompertz model demonstrates better fit than logistic model for men in age interval 85-106 years

Akaike information criterion (AIC) to compare Kannisto and Gompertz models, **women**, by birth cohort (non-Southern states)

Conclusion: In all ten cases Gompertz model demonstrates better fit than logistic model for men in age interval 85-106 years

What is the Human Mortality Database?

- The Human Mortality Database (HMD) was created to provide detailed mortality and population data to researchers, students, journalists, policy analysts, and others interested in the history of human longevity.
- **URL:** http://www.mortality.org/

The second studied dataset: U.S. cohort death rates taken from the Human Mortality Database
Selection of competing mortality models using HMD data

- Data with reasonably good quality were used: 80-106 years age interval
- Gompertz and logistic (Kannisto) models were compared
- Nonlinear weighted regression model for parameter estimates (Stata 11)
- Age-specific exposure values were used as weights (Müller et al., Biometrika, 1997)
- Model goodness-of-fit was estimated using AIC and BIC

Fitting mortality with Kannisto and Gompertz models, HMD U.S. data

Akaike information criterion (AIC) to compare Kannisto and Gompertz models, men, by birth cohort (HMD U.S. data)

Conclusion: In all ten cases Gompertz model demonstrates better fit than logistic model for men in age interval 80-106 years

Compare DMF and HMD data

Hypothesis about two-stage Gompertz model is not supported by real data
Which estimate of hazard rate is the most accurate?

Simulation study comparing several existing estimates:
- Nelson-Aalen estimate available in Stata
- Sacher estimate (Sacher, 1956)
- Gehan (pseudo-Sacher) estimate (Gehan, 1969)
- Actuarial estimate (Kimball, 1960)

Simulation study to identify the most accurate mortality indicator
- Simulate yearly \( l_x \) numbers assuming Gompertz function for hazard rate in the entire age interval and initial cohort size equal to \( 10^{11} \) individuals
- Gompertz parameters are typical for the U.S. birth cohorts: slope coefficient (alpha) = 0.08 year\(^{-1}\); \( R_0 = 0.0001 \) year\(^{-1}\)
- Focus on ages beyond 90 years
- Accuracy of various hazard rate estimates (Sacher, Gehan, and actuarial estimates) and probability of death is compared at ages 100-110

Simulation study of Gompertz mortality
Compare Sacher hazard rate estimate and probability of death in a yearly age interval
Sacher estimates practically coincide with theoretical mortality trajectory
\[ \mu_x = \frac{1}{2 \Delta x} \ln \left( \frac{l_{x+\Delta x}}{l_x} \right) \]
Probability of death values strongly underestimate mortality after age 100
\[ q_x = \frac{d_x}{l_x} \]

Simulation study of Gompertz mortality
Compare Gehan and actuarial hazard rate estimates
Gehan estimates slightly overestimate hazard rate because of its half-year shift to earlier ages
\[ \mu_x = -\ln(1 - q_x) \]
Actuarial estimates underestimate mortality after age 100
\[ \mu_x = \frac{2}{\Delta x} \left( \frac{l_{x+\Delta x}}{l_x} - 1 \right) \]

Deaths at extreme ages are not distributed uniformly over one-year interval
85-year olds 102-year olds

Accuracy of hazard rate estimates
Relative difference between theoretical and observed values, %

<table>
<thead>
<tr>
<th>Estimate</th>
<th>100 years</th>
<th>110 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of death</td>
<td>11.6%, understate</td>
<td>26.7%, understate</td>
</tr>
<tr>
<td>Sacher estimate</td>
<td>0.1%, overstate</td>
<td>0.1%, overstate</td>
</tr>
<tr>
<td>Gehan estimate</td>
<td>4.1%, overstate</td>
<td>4.1%, overstate</td>
</tr>
<tr>
<td>Actuarial estimate</td>
<td>1.0%, understate</td>
<td>4.5%, understate</td>
</tr>
</tbody>
</table>

1894 birth cohort from the Social Security Death Index
A word of caution: Data smoothing may lead to spurious mortality deceleration

What happens with simulated Gompertz law data after their Kernel smooth by Stata (default settings):

Mortality of 1894 birth cohort
Monthly and Yearly Estimates of Hazard Rates using Nelson-Aalen formula (Stata)

Sacher formula for hazard rate estimation
(Sacher, 1956; 1966)

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

- \( l_x \) - survivor function at age \( x \)
- \( \Delta x \) - age interval

Simplified version suggested by Gehan (1969):

\[
\mu_x = -\ln(1-q_x)
\]

Conclusions

- Deceleration of mortality in later life is more expressed for data with lower quality. Quality of age reporting in DMF becomes poor beyond the age of 107 years
- Below age 107 years and for data of reasonably good quality the Gompertz model fits mortality better than the logistic model (no mortality deceleration)
- Sacher estimate of hazard rate turns out to be the most accurate and most useful estimate to study mortality at advanced ages

Acknowledgments

This study was made possible thanks to:

- generous support from the
- National Institute on Aging (R01 AG028620)
- Stimulating working environment at the Center on Aging, NORC/University of Chicago
For More Information and Updates
Please Visit Our
Scientific and Educational Website
on Human Longevity:

- http://longevity-science.org

And Please Post Your Comments at
our Scientific Discussion Blog:

- http://longevity-science.blogspot.com/