# Contribution Title

Mortality Modeling

# Authors

### Tim Riffe

Research Scientist, Laboratory of Population Health, Max-Planck-Institute for Demographic Research, Rostock, Germany

Email: [riffe@demogr.mpg.de](mailto:riffe@demogr.mpg.de)

### Marília Nepomuceno

Research Scientist, Research Group on Lifespan Inequalities, Max-Planck-Institute for Demographic Research, Rostock, Germany

Email: [nepomuceno@demogr.mpg.de](mailto:nepomuceno@demogr.mpg.de)

### Ugofilippo Basellini

PhD Researcher, Institut national d’études démographiques (INED), Paris, France and Interdisciplinary Centre on Population Dynamics (CPOP), Department of Public Health, University of Southern Denmark, Odense, Denmark

Email: [ugofilippo.basellini@ined.fr](mailto:ugofilippo.basellini@ined.fr)

# Synonyms

Mortality Laws

# Definition

Mortality models approximate mortality patterns or dynamics over age and time. An age pattern of mortality can be any mathematical function of mortality, such as rates, probabilities, survivorship, or death distributions. Such functions may be modeled in the form of a life table or a simplified model of parameters. Mortality models in general fall into three main categories: (i) models designed to help understand regularities in mortality patterns and dynamics, for example where population-level mortality patterns are modeled as an emergent property of dynamics at the individual level, (ii) those that aim to predict mortality patterns, for example for purposes of pension provisions, and (iii) those aimed at mortality measurement for purposes of mortality and health monitoring. In the following, mortality modeling refers to models of mortality measurement at the population level.

# Overview

Mortality modeling aims to measure, approximate, or predict mortality patterns and changes over age and time. The first to explicitly model mortality was Graunt (1662), who summarized mortality conditions in London in the first application of an empirical life table. Life tables use a set of mathematical identities to translate present mortality conditions into their long run consequences, which makes them models (crossref life tables). The number of data input elements in the life table can be thought of as its parameters, and life table has about as many parameters as it has age classes. Simple mortality patterns that occur in limited age ranges, can be reproduced by more parsimonious models based on fewer parameters, and this is what is usually thought of as parametric mortality modeling.

Parametric mortality models are usually designed to describe the so-called force of mortality (link to entry), or rate of attrition from death. From these models, one can derive several more substantively interpretable mortality functions, such as the remaining life expectancy or the number of survivors at each age. Thus, the ability of mortality models to accurately reflect the patterns and dynamics of mortality has great implications in measures widely used for monitoring and understanding health. For example, life expectancy at birth summarizes the risk of death over the entire age range, and its accuracy is therefore to some extent dependant on model performance in all ages. Modeling choices, such as whether to include a component for young adult excess mortality, or decisions on whether and how to extrapolate mortality in the oldest ages can have notable impacts on important summary measures such as life expectancy.

The first parametric mortality model to be proposed was that of Gompertz (1825), which describes mortality in adult ages very well using just two parameters, and is still widely used today (see section XXX). Makeham (1860) added a third parameter for background mortality, and this improved the model fit. Focus on accuracy in adult and older ages in these early years was motivated by life insurance and annuities, and still is today. A decade later, Thiele (1871), proposed a model of seven parameters which fit the contours of mortality over the entire age range. Unprecedented mortality improvements in the 20th century and other unique changes in the age pattern of mortality (young adult excess mortality, the HIV crisis) have motivated the development of more complex models of mortality or different modeling approaches.

The last century recorded remarkable changes in patterns of health and disease. The leading causes of death shift from infections and external causes to degenerative diseases (Omran 1971). Life expectancy improvements driven by reduced mortality among children and young adults became increasingly driven by delays in death at older ages. These qualitative statements of mortality change constitute a model insofar as they are anticipated in populations where no such shift has been observed. Sustained improvements in survival to older ages in most industrialized countries increased the need for accurate measurement of old-age mortality and interest in measuring and modeling mortality beyond age 100 (centenarians) and 110 (supercentenarians)(link to entry). Due to the very small number of observations of centenarians and supercentenarians, and the scarce availability of accurate data for these individuals, modeling mortality at the oldest ages presents unique challenges. Aggregate mortality in these advanced ages is usually modeled with either exponential, logistic, or an intermediate pattern (Gompertz 1825; Perks 1932; Beard 1971; Kannisto 1994; Thatcher et al. 1998).

Another important approach to modeling in contexts of sparse or incomplete data, is to apply one of the so-called model life table systems (link to entry?). Model life tables consist in two parts: (i) A standard mortality pattern that covers the full age range, potentially a blend of many higher quality life tables and (ii) an approach to adjust the standard to match a reduced set of measured values (see section XXX for an application). Often the measured values (the model parameters) are measures in two or more ages. Usually one refers to a system An intuitive introduction to model life table approaches can be found in Moultrie et al. (2013).

Modeling choices also play a key role in mortality forecasting. Most methods of mortality forecasting extrapolate both age patterns and trends over time. Extrapolative methods may also use summary mortality measures, such as life expectancy to guide or constrain the forecast. Whether based on summary indices or age-specific mortality measures, extrapolative methods employ a mortality model, as described by a life table or a mathematical function. For example, the Lee-Carter method (1992) is a widely used extrapolative method to forecast mortality over age and time (crossref to LC).

Improvements in mortality modeling techniques lead to better measurement and predictions of mortality, which leads to more accurate aggregate mortality measures, more complete measurement in contexts of sparse or incomplete data, more plausible projections, and better understanding of the mechanisms behind changes in the age-pattern of mortality.

# Key Research Findings

This section contains an overview of some key findings in mortality modeling over the last centuries.

### Gompertz law

Some findings are so empirically regular that they are referred to as *laws*, although when applied in models they are assumptions. The Gompertz (1825) law states that the force of mortality increases exponentially over adult ages, and it is the first and most important of the mortality laws. This assumption has proven to be an accurate and useful representation of the mortality pattern among diverse populations throughout the last two centuries. As such, the Gompertz law of mortality is today one of the most well-known models of human mortality, and it is still widely employed in demographic and actuarial analyses. Fig. 1 shows the fit of the Gompertz model to the male population of Switzerland in 2010 based on data from the Human Mortality Database (2018).The graph shows the assumption of exponential (linear in log scale) increase of mortality with age to be a good representation of the human mortality pattern.

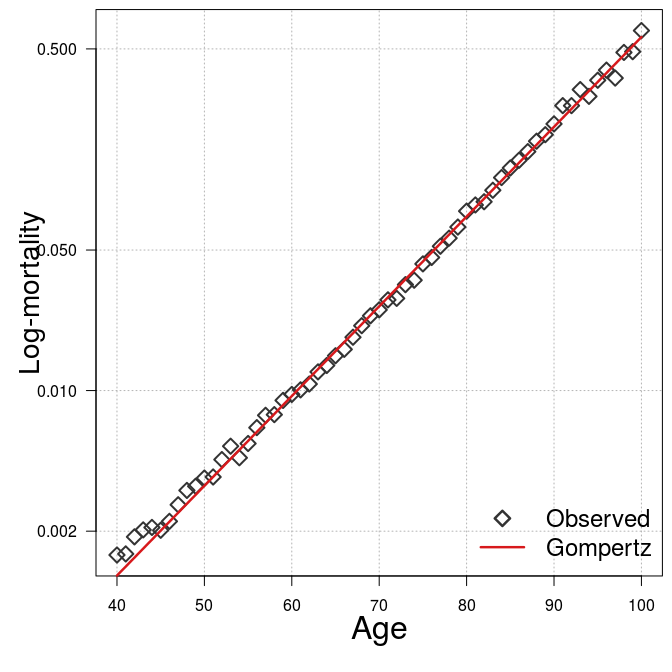


Figure 1 Observed and fitted death rates using a log-linear (Gompertz) model for males aged 40 to 100 in Switzerland in the year 2010.

### Mortality over the life course

In 1871, the Danish astronomer and actuary Thorvald Thiele introduced the a partition of human mortality pattern into three groups that operate principally on childhood, middle, and older ages, respectively. These three components of mortality represent the risk of death at different times in life, and they have been observed to follow a stable pattern throughout history. The first component, for infant and child mortality, decreases steadily with age, with a sharp initial reduction after birth. The second component, known as the young adult *hump* for its convex shape, describes mortality after the onset of puberty and stretching into younger adult ages. Specifically, the component is often due to excess mortality from external and behavioral deaths, especially among males, but also historically from tuberculosis for both sexes and maternal mortality for females. The third component is generally denoted as senescent mortality, which was Gompertz’s centre of attention. As noted above, the component starts at middle-adult ages and increases exponentially with age. Several parametric (for example Heligman and Pollard 1980) and non-parametric approaches have been proposed to decompose the human mortality pattern in a similar way. Fig. 2 shows the partition obtained by the non-parametric model of Camarda et al. (2016), applied to the same population of Fig. 1. The aggregate mortality pattern divides clearly into these three substantively interpretable components of mortality over the life course.

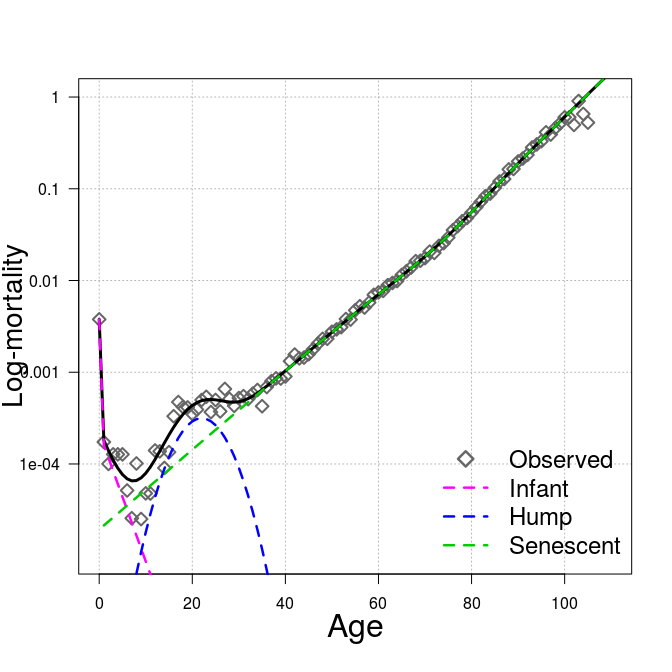


Figure 2 Additive partition of observed death rates into infant, hump, and senescent components for males in Switzerland in the year 2010 (data: HMD).

### Debate on the plateau

There is widespread agreement that mortality rates increase exponentially between mid adulthood through ages 80 to 90. However, there is no consensus regarding the mortality rate trajectory at the most advanced ages. Some studies conclude that the exponential growth of mortality with age is followed by a period of deceleration, with slower rates of mortality increase at the oldest ages leading to a plateau of human mortality (Barbi et al. (2018), Gampe 2010; Robine and Vaupel 2001; Robine et al. 2005). Models including Kannisto (1994), Beard (1971) and Perks (1932) produce a mortality plateau. Others conclude that the mortality deceleration is an artifact of data with lower quality, and that mortality continues to grow exponentially through the highest ages (Gavrilov and Gavrilova 2011, 2019; 2014; Newman 2018).

### Rectangularization

Mortality rates can be translated to a so-called survivor curve using lifetable techniques. The survivor curve represents the fraction of a cohort that would survive from birth until each age if they were exposed over the life course to the death rates of a given life table. Rectangularization is defined by the process during which the shape of the survival curve approaches the shape of a rectangle due to reductions in premature mortality and the concentration of the deaths in a narrow range of older ages. Fig. 3 shows an example of rectangularization in Swedish male survival curves from 1900 to 2010 (HMD). Fries (1980) argued that rectangularization occurs when life expectancy at birth approaches the upper limit of human lifespan due to a decrease in variability in age-at-death. In other words, when lifespan variability decreases, deaths are compressed at older ages. Since the second half of the twentieth century, lifespan variability has been stagnating and life expectancy continued to increase in high-income countries, resulting in a process known as mortality shifting (Kannisto 2000; Bongaarts 2005; Bergeron-Boucher et al. 2015). Mortality shifting occurs when life expectancy increases due to a shift in the death distribution toward older ages with nearly constant lifespan variability (Vaupel 1986; Yashin et al. 2001; Bongaarts 2005; Canudas-Romo 2008).

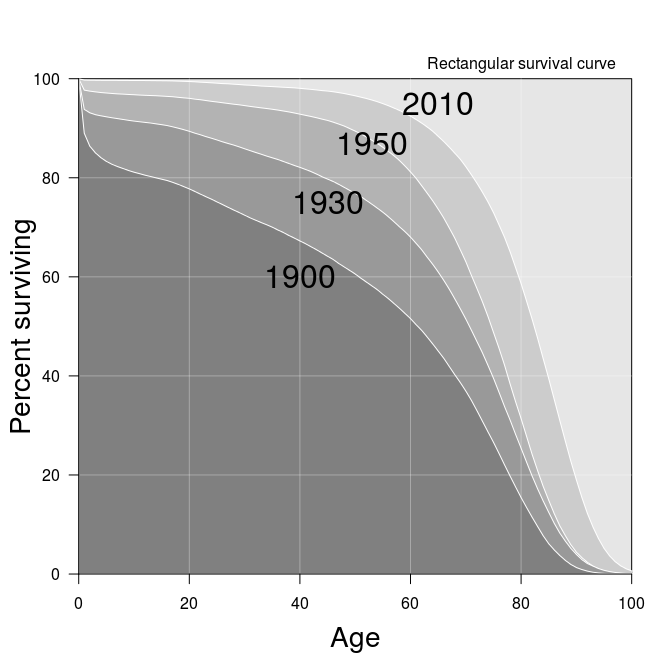


Figure 3 Survival curves of Swedish males in the years 1900, 1930, 1950, and 2010 appear to approach a rectangular shape (data: HMD).

### Oeppen-Vaupel line

Oeppen and Vaupel (2002) pointed out a striking regularity in human mortality since 1840: The highest level of observed life expectancy each year has been rising at a steady pace of almost 3 months per year over the entire period. The authors analyzed the evolution of female best-practice life expectancy at birth (that is, the life expectancy of the country holding the highest level in the whole world in a given calendar year) from 1840 to 2000, finding a linear trend that increased at a steady pace of almost 3 months per year during the last 160 years. Fig. 4 shows this finding, which has had direct implications in a number of modeling domains, especially mortality forecasting (Torri and Vaupel 2012; Pascariu et al. 2018).

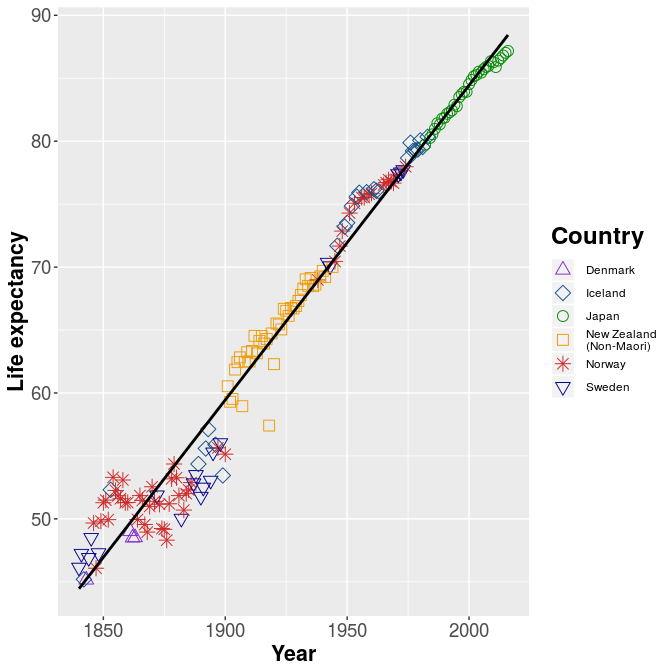


Figure 4 Best-practice female life expectancy at birth from 1840 to 2016 (data: HMD).

# Examples of Application

### Model life tables

Model life tables comprise a family of techniques designed to relate partial mortality information to a full age pattern of mortality, which can then be used to calculate standard mortality indicators such as life expectancy. Many variants exist and have long been in use, especially in the context of mortality measurement in countries without complete vital registration (United Nations 1982; Coale et al. 1983; United Nations Population Division 2012; Wilmoth et al. 2012; Clark 2015). Model patterns are derived by reducing the mortality patterns in a large collection of life tables into a smaller set of standard relationships. A given standard is then scaled and warped according to a small number of directly measureable and usually intuitive parameters.

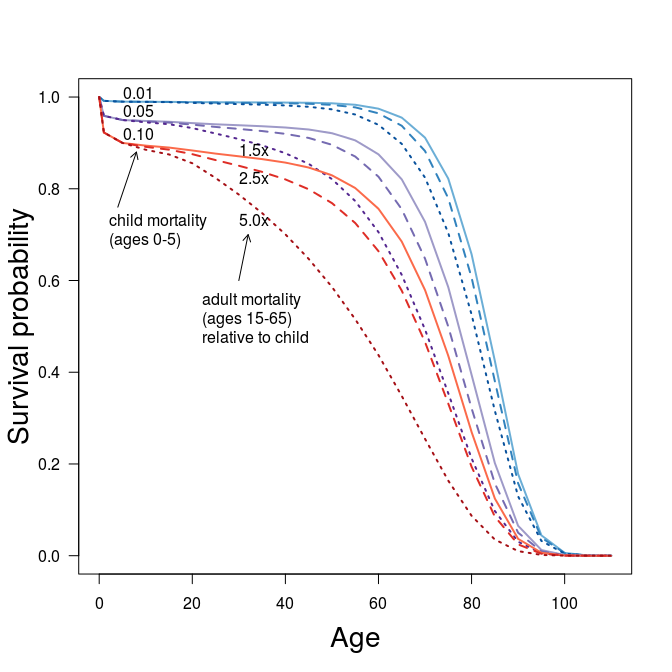


Figure 5 A demonstration of some of the flexibility given in the Wilmoth et al (2012) model life table approach. A standard is derived from HMD male life tables. From this, nine different survivor curves are derived from combinations of specified child and adult mortality. Three levels of child mortality (0.01,0.05,0.10) are paired with three levels of relative adult mortality, 1.5, 2.5, and 5.0 times higher than the given child mortality estimate, respectively.

Fig. 5 demonstrates the flexibility that most model systems can achieve, in this case using the Wilmoth et al. (2012), where the standard mortality pattern is derived from male life tables in the Human Mortality Database (2018). Results are calculated using the MortalityEstimate R package (Pascariu et al. 2019). In practice the standard could be derived from any corpus of life tables. For instance, a collection of lifetables in populations under different levels of HIV prevalence would give a different standard pattern. The standard is then adjusted according to specified estimates of child and adult mortality. Child mortality is defined as the probability of not reaching age 5, and adult mortality is here defined as the probability of a 15-year-old not reaching age 65. These are just two of several potential parameterizations of this model. These combinations reveal a wide range of realistic mortality patterns that can be derived from just two parameters that can be estimated independently.

For the case of gerontological research, model life tables such as that demonstrated can 1) extend the geographical range of research to populations not usually included in research due to data limitations, 2) extend estimation to subpopulations for which complete life tables cannot be directly calculated, 3) systematize the relationship between mortality in different age groups to infer older age mortality from younger ages when older ages are not directly observed (compare with extrapolation using parametric mortality laws).

### Extrapolation and smoothing

The Human Mortality Database (HMD) uses a parametric procedure to smooth death rates above age 80 in life tables. Specifically, the Kannisto (1994) model of old-age mortality is used to remove the randomness inherent to the observed rates at older ages, thus obtaining ``an improved representation of the underlying mortality conditions’’ (Wilmoth et al. 2017, p. 34). The model is fitted to ages 80 and above, separately for males and females, assuming a Poisson distribution for death counts. The estimated parameters yield a smooth function of death rates over age, which is used to replace the observed rates for all ages above a variable age . In particular, is defined as the lowest age where there are at most 100 deaths (male or female), but it is constrained to fall within the age interval [80, 95] (for additional information, see Wilmoth et al. (2017)).

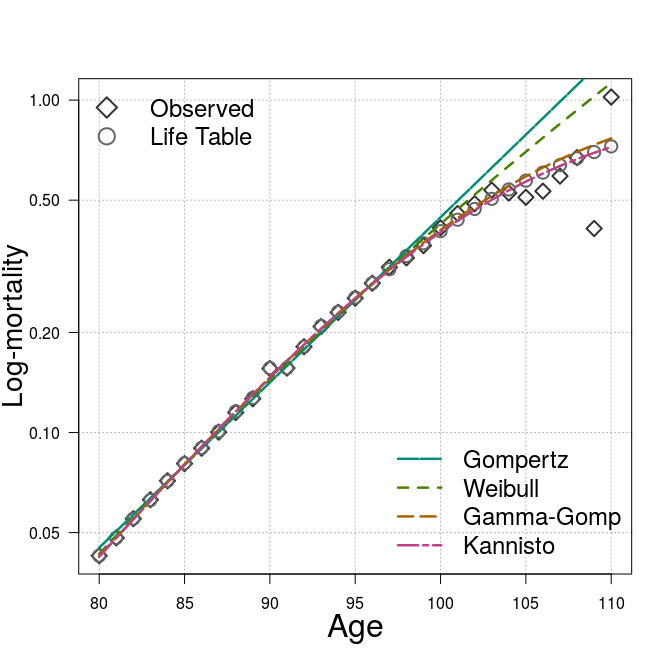


Figure 6 Observed and life table death rates for females in England and Wales in the year 2010, with fit of four different parametric mortality models. The Gompertz model is log linear, whereas the Kannisto and Gamma-Gompertz models deviate after age 95, tapering toward a plateau. The Weibull extrapolation is intermediate in this case.

Fig. 6 shows an example of the HMD smoothing procedure compared with other selected model options (models were fit using the *MortalityLaws* R package Pascariu 2018) as well as the original observed death rates for females in England and Wales in 2010 (HMD). HMD life table rates are identical to observed rates below age 95, which is age for this population. In particular, the life table death rates are derived from the Kannisto mortality model, as discussed above. The HMD Kannisto fit differs from the In addition, three other well-known models of mortality to this population: the Gompertz, the Weibull and the Gamma-Gompertz model. The fitted lines show that the Gamma-Gompertz model produces estimates that are very close to the Kannisto model, while the Gompertz and Weibull model overestimate the observed pattern of mortality at the highest ages.

# Future Directions of Research

Future research in mortality modeling will extend the applications of presently available modeling techniques to yet-understudied populations, either due to innovations to increase applicability or due to new data becoming available. Several thematic areas for future research are evident:

### Limits to life and the mortality plateau

The scientific community is far from reaching a consensus on whether and in what way the human longevity is subject to fundamental limits, either in terms of the maximum age attained or a maximum death rate (compare Olshansky et al. (1990) and Barbi et al. (2018)) , and this is a fundamental concern for the field of gerontology. For example, approaches based on the consideration of physiological constraints have not been harmonized with demographic approaches. Data collection and validation of centenarians and super-centenarians is ongoing, and the number of people reaching these ages worldwide has been increasing rapidly (Vaupel 2010). In the coming decades the increased amount of data on the longest lived will be able to provide more nuance to the question of whether or not and at what level mortality levels off in old age, and ideally lead to a consensus between disciplines on the character of mortality among the extremely longevous.

### New population definitions

Many methods to harmonize mortality estimation from diverse sources and to make deficient or incomplete vital register data more usable in standard life table applications [XXXXX UB comment: isn’t there a missing verb here? Like “may methods AIM to …”]. The same push to improve mortality estimation worldwide is also relevant to populations with good data, which can benefit from the same tools in order to differentiate mortality outcomes for small, highly local, difficult-to-observe, or partially observed subpopulations. Large populations may be further stratified in creative ways (for example, by different definitions of income, wealth, and capital) that might reveal group differences in mortality patterns and levels. Likewise, new registers will better identify health risk factors that differentiate mortality outcomes.

### Coherent modeling

Innovation is needed to overcome modeling difficulties arising from the two-sided compositional nature of mortality: First, mortality is measured with respect to some definition of population at risk, which itself is necessarily a composition of heterogeneous risk levels. A standard approach to model the unobserved risk structure of mortality would benefit the field greatly. Second, mortality is a composition of outcomes, such as causes of death. This raises difficulties in practice, for example in projecting mortality by cause of death, or when smoothing mortality jointly by cause of death. Such operations would also benefit from standard solutions.

### Lifespan inequality

Most mortality research has been focused on life expectancy, or average length of life as the primary life table outcome, but recent research highlights the necessity of estimating and monitoring lifespan uncertainty as a primary indicator of population health status, and as a fundamental kind of population inequality (Raalte et al. 2018). Mortality modeling often relates parameters to life expectancy, but not lifespan inequality, and lifespan inequality is rarely used to calibrate models (Bohk-Ewald et al. 2017). Future research will measure social differences in lifespan inequality, and modeling will either account for inequality or be parameterized in terms of inequality.

### Model translation

To the extent that mortality modeling has bifurcated into subfields for the purposes of measurement, prediction, and explanation, there is potential for innovation by translating models to other ends. For example, Sharrow and Anderson (2016) offer a process-based model to partition mortality into intrinsic and extrinsic components, and Camarda et al (2016) propose a decomposition of mortality into three substantive components for infant, young adult excess, and old-age mortality. Both are in the first place a fundamental question of measurement, but hold the potential to offer new approaches to prediction or relate to models of understanding. Many instances of model cross-fertilization of this kind are likely to be proposed in future research.

# Summary

# Cross references

# References

Barbi E, Lagona F, Marsili M et al (2018) The plateau of human mortality: Demography of longevity pioneers. Science 360:1459–1461

Beard R (1971) Biological aspects of demography. In: Brass W (ed). Taylor & Francis Ltd, pp 57–68

Bergeron-Boucher M-P, Ebeling M, Canudas-Romo V (2015) Decomposing changes in life expectancy: Compression versus shifting mortality. Demographic Research 33:391–424

Bohk-Ewald C, Ebeling M, Rau R (2017) Lifespan disparity as an additional indicator for evaluating mortality forecasts. Demography 54:1559–1577

Bongaarts J (2005) Long-range trends in adult mortality: Models and projection methods. Demography 42:23–49

Camarda CG, Eilers PH, Gampe J (2016) Sums of smooth exponentials to decompose complex series of counts. Statistical Modelling 16:279–296

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

Clark SJ (2015) A singular value decomposition-based factorization and parsimonious component model of demographic quantities correlated by age: Predicting complete demographic age schedules with few parameters. arXiv preprint. doi: [arXiv:1504.02057](https://doi.org/arXiv:1504.02057)

Coale A, Demeny P, Vaughan B (1983) Regional model life tables and stable populations. New York NY/London England Academic Press 1983.

Fries J (1980) Aging, natural death, and the compression of morbidity. The New England journal of medicine 303:130–135

Gavrilov LA, Gavrilova NS (2011) Mortality measurement at advanced ages: A study of the social security administration death master file. North American actuarial journal 15:432–447

Gavrilov LA, Gavrilova NS (2019) Late-life mortality is underestimated because of data errors. PLoS biology 17:e3000148

Gavrilova NS, Gavrilov LA (2014) Biodemography of old-age mortality in humans and rodents. Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences 70:1–9

Gompertz B (1825) XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to francis baily, esq. FRS &c. Philosophical transactions of the Royal Society of London 115:513–583

Graunt J (1662) Natural and Political Observations Made upon the Bills of Mortality. London. American edition edited by Walter F. Willcox, John Hopkins Press, Baltimore, 1939.

Heligman L, Pollard JH (1980) The age pattern of mortality. Journal of the Institute of Actuaries 107:49–80

Human Mortality Database (2018) University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany)

Kannisto V (2000) Measuring the compression of mortality. DEMOGRAPHIC RESEARCH 3:24

Kannisto V (1994) Development of oldest-old mortality, 1950-1990: Evidence from 28 developed countries. University Press of Southern Denmark

Lee RD, Carter LR (1992) Modeling and forecasting US mortality. Journal of the American Statistical Association 87:659–671

Makeham WM (1860) On the law of mortality and the construction of annuity tables. The Assurance Magazine, and Journal of the Institute of Actuaries 8:301–310

Moultrie TA, Dorrington R, Hill A et al (2013) Tools for demographic estimation. International Union for the Scientific Study of Population

Newman SJ (2018) Errors as a primary cause of late-life mortality deceleration and plateaus. PLoS biology 16:e2006776

Oeppen J, Vaupel J (2002) Demography. Broken limits to life expectancy. Science (New York, NY) 296:1029–1031

Olshansky SJ, Carnes BA, Cassel C (1990) In search of methuselah: Estimating the upper limits to human longevity. Science 250:634–640

Omran AR (1971) The theory of epidemiological transition. Milbank Memorial Fund Quarterly 49:509–538

Pascariu MD (2018) MortalityLaws: Parametric mortality models, life tables and hmd

Pascariu MD, Aburto JM, Basellini U, Canudas-Romo V (2019) MortalityEstimate: Indirect estimation methods for measurement of demographic indicators

Pascariu MD, Canudas-Romo V, Vaupel JW (2018) The double-gap life expectancy forecasting model. Insurance: Mathematics and Economics 78:339–350. doi: [10.1016/j.insmatheco.2017.09.011](https://doi.org/10.1016/j.insmatheco.2017.09.011)

Perks W (1932) On some experiments on the graduation of mortality statistics. Journal of the Institute of Actuaries 63:12–40

Raalte AA van, Sasson I, Martikainen P (2018) The case for monitoring life-span inequality. Science 362:1002–1004

Sharrow DJ, Anderson JJ (2016) Quantifying intrinsic and extrinsic contributions to human longevity: Application of a two-process vitality model to the human mortality database. Demography 53:2105–2119

Thatcher AR, Kannisto V, Vaupel JW (1998) The force of mortality at ages 80 to 120. Odense University Press

Thiele TN (1871) On a mathematical formula to express the rate of mortality throughout the whole of life, tested by a series of observations made use of by the danish life insurance company of 1871. Journal of the Institute of Actuaries 16:313–329

Torri T, Vaupel JW (2012) Forecasting life expectancy in an international context. International Journal of Forecasting 28:519–531

United Nations (1982) Model life tables for developing countries. United Nations

United Nations Population Division (2012) World population prospects 2012: Extended model life tables. United Nations, New York

Vaupel JW (1986) How change in age-specific mortality affects life expectancy. Population Studies 40:147–157

Vaupel JW (2010) Biodemography of human ageing. Nature 464:536

Wilmoth J, Andreev K, Jdanov D et al (2017) Methods protocol for the human mortality database. University of California, Berkeley and Max-Planck-Institute for Demographic Research

Wilmoth J, Zureick S, Canudas-Romo V et al (2012) A flexible two-dimensional mortality model for use in indirect estimation. Population studies 66:1–28

Yashin AI, Begun AS, Boiko SI et al (2001) The new trends in survival improvement require a revision of traditional gerontological concepts. Experimental Gerontology 37:157–167