# 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 or 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 forecasting, 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 (in progress)

Gerontologists, actuaries, policy makers and public health practitioners are interested in the number of years to be lived by individuals (life expectancy) and or in the number of survivors at different ages. These measures are widely used to monitor health, promote the efficiency of health services, and calculate annuities. Life expectancy and the number of survivors are mortality functions derived from an age pattern of mortality. This is a reason why since the 17th century researchers have been devoted to developing mortality models.

Graunt (1662) and Halley (1693) were the first who explicitly modeled mortality through empirical life tables. But only in 1725, the first theoretical mortality model was published by DeMoivre. A century later, a well-established age pattern of mortality was developed by Gompertz. The author showed that mortality increases exponentially with age from adulthood until ages around 80-90. In 1867, Makeham modified the Gompertz model by including a parameter to capture the background mortality in order to better describe adult mortality. Since then, to reproduce reality, mortality models have been accompanying changes in the age pattern over time. Moreover, they attempt to describe mortality from childhood to older ages. To achieve that, mortality models had become more complex. For example, Heligman and Pollard (1980) proposed a model to describe mortality for the entire age range by decomposing the age pattern of mortality into child mortality, adult mortality which can describe the accident hump, and old-age mortality.

The aging of populations increases the need for accurate measurement of mortality at advanced ages (above age 80). However, due to the very small number of survivors at the most advanced ages, and the scarce availability of accurate data for these individuals, modeled mortality at older ages is challenging. To date, logistic and exponential models has been used to describe the age pattern of mortality at very high ages. In 1932, Perks was the first who proposed a logistic function to describe adult mortality. Since then, Beard (1971), Kannisto (1994), Thatcher et al. (1998) also suggested a slower rate of mortality increase at very old ages.

Measurement changes in age patterns of mortality are one of the main motivation for modeling; however, there are at least three motivations for modeling, which correspond to meta approaches: measurement, explaining and predicting.

### Measurement (reduce this!)

More often the business of research is to measure the mortality (longevity) of (stratified) populations, which except in the most privileged data environments, always invokes some degree of modeling. For example, even if the researcher is satisfied with relative risks, in many cases the researcher will often measure this using a proportional hazards assumption, possibly with a log linear mortality under the hood. In the more general case, where stratified lifetables are the objective (e.g to compare expectancies or lifespan disparities), modeling for measurement might rely on the application of standard mortality patterns (for example model lifetable look-ups, Brass relational logit fitting of an arbitrary standard, or the Wilmoth log quad relational method to an arbitrary standard), which usually proceeds by specification of 1, 2, or 3 substantive parameters, like life expectancy, infant or child mortality, or adult mortality in a fixed age range (possibly all originating from surveys) to select or warp the given standard. In this case the standard and its parameter mapping constitute the mortality model. In the same sense parametric models, such as the Gompertz or Kannisto models are often used to smooth or extrapolate mortality in old ages for purposes of noise (or age heaping) reduction or lifetable closeout, and this is an example of mortality models applied to measurement. Even stable population theory is applied to measurement in the case of the generalized growth balance adjustment for mortality under/over registration (relative to population registration)- in this case the assumptions invoked are the model. In the area of statistical inference for measurement mortality is modeled via the estimator family. For example, a Poisson counting process assumption for deaths given a certain level of exposure is another commonly used model, but others, e.g. log normal rate distributions arise from time to time. Here again the objective is measurement. The same is true of smoothing in general, where the objective is to parse data down to its signal, where the signal is the measurement.

### Explaining (reduce this!)

Since mortality rates are an emergent property of individuals (individuals do not have rates, but group rates are an instrument to approximate the unmeasureable notion of individual risk), one branch of mortality modeling tries to produce process models at the individual level, such that empirical properties of human mortality arise in the aggregate (see Finkelstein; Anatoli; Anderson & Li, Vaupel, Alan Cohen). Such models linking individual senescence to aggregate mortality might take into account concepts such as shocks to the organism that deplete some store of vitality or frailty, or aspects of individual resuscitation, and they might be deterministic or stochastic. This class of model seeks to answer the question of why mortality has the character that it does, or else to infer something about the unobserved distribution of frailty or the experience of vitality and shocks within individuals. This kind of modeling is premised to a certain extent of human mortality having some sort of permanent character. Secular changes in the character of mortality (e.g. rectangularization) are often taken into account with ex post model generalizations, and they have thus far not turned out to be ex ante predictions of these models. For this reason, I say that this (extended) family of mortality modeling aims to explain, and not to measure or predict mortality.

### Predicting (reduce this!)

Mortality prediction might refer to projection, retrojection, or even interpolation for missing observations. In any of these cases, the purpose for model abstraction is to reduce the number of parameters to project to those that i) are able to reproduce the age pattern, and ii) themselves have evolved in some regular way over time, for example linearly. For example, the Lee-Carter forecasting model takes an array of mortality rates by age and time (a \* t initial parameters to summarize) and reduces it to (approximately) 2 \* a + 2 parameters, of which all but one are modeled as fixed in time, and a single parameter is forecast, often but not necessarily with a time series model. While LC is a particularly popular prediction model for projection, retrojection, and interpolation (with many variants), all prediction models bear some resemblance in that usually one (or sometimes two or more correlated parameters) are extrapolated in a first step, and then the mortality age pattern is invoked via the specified mapping (which is hopefully time invariant). Since this branch of mortality modeling is not concerned with explaining, the components of such models are not bound to having substantive interpretations, and sometimes process-agnostic matrix decomposition methods (SVD, PCA, or others) are used to reduce the parameter space. Indeed, these same tactics are also used for the last (but not least!) objective of measurement.

# Key Research Findings

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

### Gompertz law

One of the most important contributions to modeling mortality traces back to almost two centuries ago. The British mathematician and actuary Benjamin Gompertz (1825) theorized that the adult age-pattern of the force of mortality, specifically around the age range 30-90, increases exponentially with age. This assumption has proven to be a pretty accurate representation of the mortality pattern of several 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. The first panel of Fig 1 shows the fit of the Gompertz model to the male population of Switzerland in 2010. As the graph shows, the assumption of exponential (linear in log scale) increase of mortality with age is still a valid representation of the human mortality pattern.

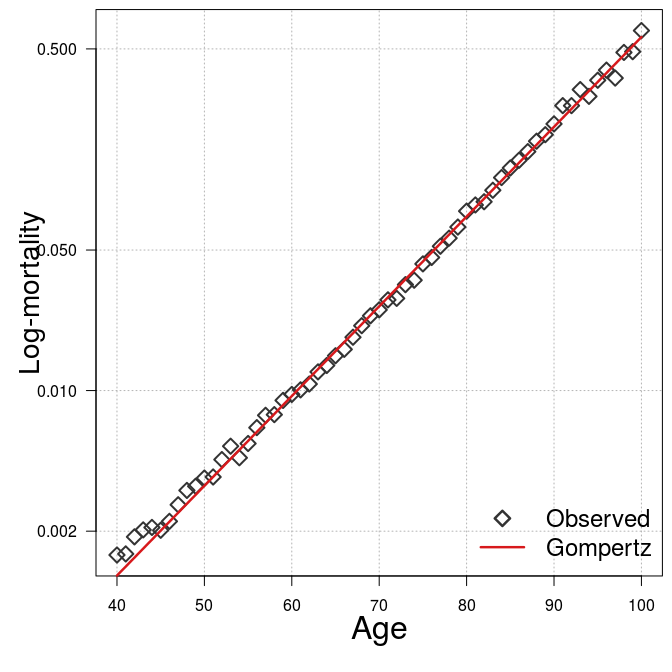


Figure 1 Observed and fitted death rates using a log-linear (Gompertz) model for males aged 40 to 100 in Switzerland in the year 2010.

### Negative senescence and hump

Another key finding regarding the age-pattern of mortality was first described a few years after Gompertz. In 1871, the Danish astronomer and actuary Thorvald Thiele introduced the decomposition of the human mortality pattern into three groups that operate principally, or almost exclusively, upon childhood, middle, and old 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 represents infant and child mortality, which decreases steadily with age, with a sharp reduction after birth. The second component, known as the accident hump for its particular shape, describes mortality after the onset of puberty and stretching into younger adult ages. Specifically, the component captures the accident deaths due to the risk-taking behavior of young men and, historically, mortality related to childbearing in women. Finally, 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 and non-parametric approaches have been proposed in the last few decades to decompose the human mortality pattern. The second panel of Fig. 2 shows this decomposition, applied to the same population as above. The three components of mortality clearly emerge from the graph.

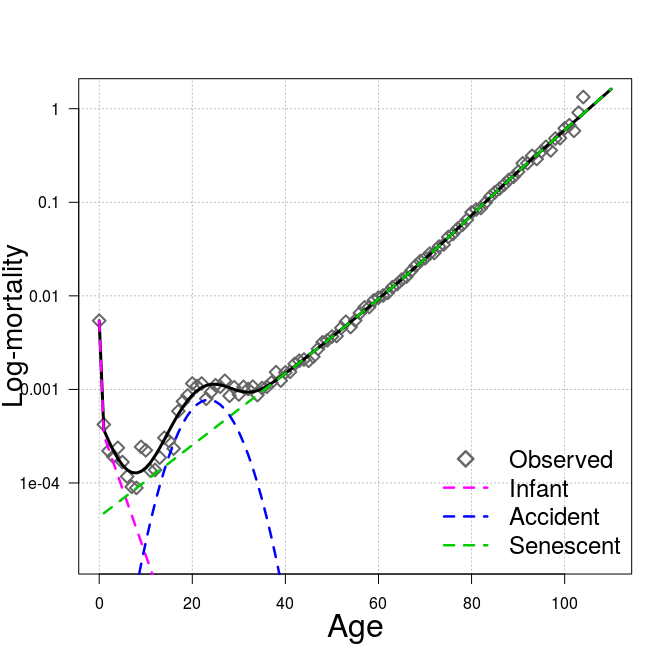


Figure 2 decomposition of the observed death rates into infant, young adult excess, and senescent components for males in Switzerland in the year 2010.

### Debate on plateau

There is a general agreement that mortality increases exponentially from mid-adult to ages 80-90, as described by the Gompertz law. However, there is not yet a consensus regarding the mortality trajectory at the most advanced ages. Some studies suggested 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, creating a plateau of human mortality (Barbi et al. (2018), Gampe 2010; Robine and Vaupel 2001; Robine et al. 2005). Another group of researchers claimed that the mortality deceleration in later life or the mortality plateau is more expressed for data with lower quality, and hence mortality continues to grow exponentially at the highest ages (Gavirilov and Gavirilova 2011, 2015, 2019). Logistic curves like Kannisto (Kannisto 1994), Beard (1971) and Perks (Perks 1932) model the mortality plateau, while the Gompertz law rises exponentially with age.

### Rectangularization

Rectangularization is defined by the process when the shape of the survival curve becomes more rectangular due to reductions in premature mortality and the concentration of deaths at older ages. In 1980, James Fries claimed that the rectangularization occurs when life expectancy at birth approaches the average lifespan due to a decrease in variability in age-at-death (Fries 1980). In other words, when lifespan variability decreases, deaths are compressed at older ages. However, 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 by mortality shifting (Kannisto 1996; Boongarts 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).

### Oeppen-Vaupel line

Oeppen and Vaupel (2002) showed one of the most striking regularities observed in human mortality during the last centuries: the highest level of observed life expectancy has been rising at a steady pace of almost 3 months per year during the last 160 years. The authors analyzed the evolution of the 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 that the trend is very linear, rising at a steady pace of almost 3 months per year during the last 160 years. Fig. 3 shows this remarkable finding, which has had direct implications in a number of spheres, for example, mortality forecasting. The best-practice line has indeed been employed in some forecasting methodologies (cite).

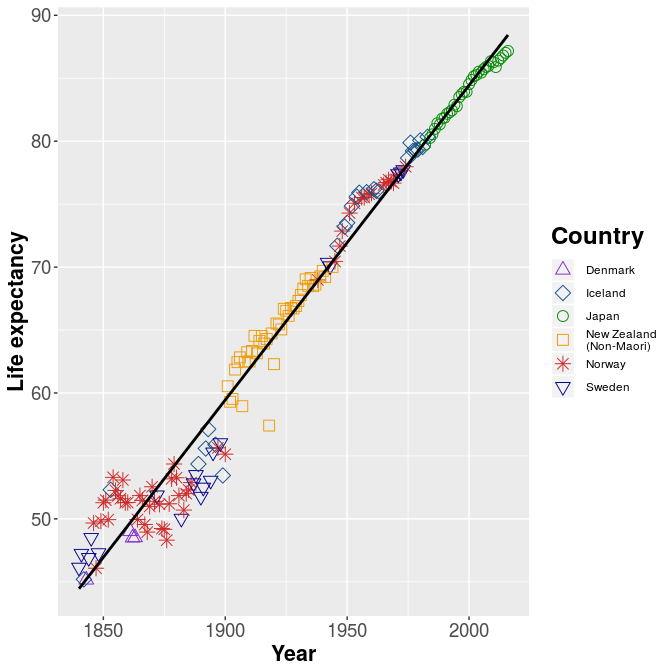


Figure 3 Best-practice female life expectancy at birth from 1840 to 2016. data: HMD.

# Examples of Application

### Model lifetables

Model lifetables 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 lifetables into a smaller set of relationships that when scaled and warped according to a small number of (usually) intuitive parameters.

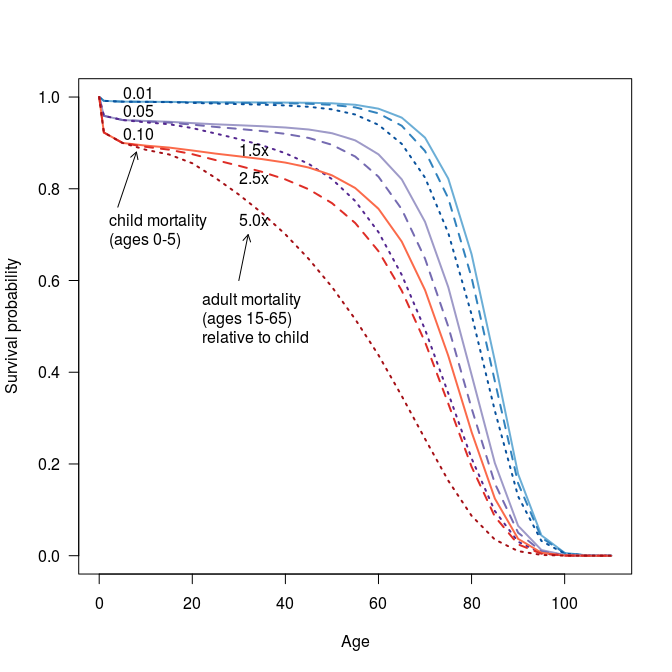


Figure 4 A demonstration of some of the flexibility given in the Wilmoth et al (2012) model lifetable approach. A standard is derived from HMD male lifetables. From this, nine different survivorships are derived from combinations of specified child and adult mortality. Three levels of child mortality (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. 4 demonstrates the flexibility that most model systems can achieve, in this case using the Wilmoth et al approach (2012), where the standard mortality pattern is derived from male lifetables 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 lifetables. 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 lifetables 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 lifetables 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 old-age death rates in period life tables. Specifically, the Kannisto model of old-age mortality (Thatcher et al. 1998) 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 yields smooth death rates, which are used to replace the observed rates for all ages above 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)).

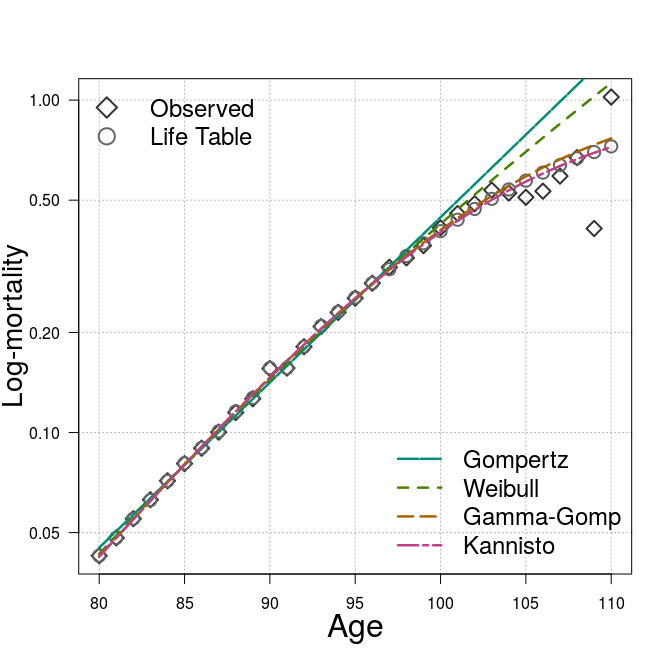


Figure 5 Observed and lifetable death rates for females in England and Wales in the year 2010, with fit of four different parametric mortality models. The Gompertz model is log linear, whereas the Kannisto and Gamma-Gompertz models deviate after age 95, tapering toward a plateau. The Weibull extrapolation is intermediate in this case.

The right panel of Fig. 5 shows an example of this procedure. This figure shows the observed and lifetable death rates for females in England and Wales in 2010 (HMD). The two rates depart from each other after age 95 (which is age for this population, as less than 100 deaths were observed from the age-group 106). In particular, the life table death rates are derived from the Kannisto mortality model, as discussed above. 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 lifetable applications. 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 lifetable 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

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

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

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.

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

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

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

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

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

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 (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