DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 51, ARTICLE 18, PAGES 595–624
PUBLISHED 6 SEPTEMBER 2024
http://www.demographic-research.org/Volumes/Vol51/18/
DOI: 10.4054/DemRes.2024.51.18

Research Article

Makeham mortality models as mixtures: Advancing mortality estimations through competing risks frameworks

Silvio C. Patricio

Trifon I. Missov

Contents

Demographic Research: Volume 51, Article 18
Research Article

Makeham mortality models as mixtures: Advancing mortality estimations through competing risks frameworks

Silvio C. Patricio[1]

Trifon I. Missov[2]

Abstract

BACKGROUND
The Makeham term serves as a fundamental component in mortality modeling, offering a constant additive hazard that accounts for background mortality factors usually unrelated to the aging process. This term, widely employed in mortality analysis, provides a crucial mechanism for capturing mortality risks unrelated to age-related deterioration.

OBJECTIVE
The objective of this paper is to explore how Makeham models, which are widely used for studying mortality, can be understood and analyzed within the context of competing risks. The paper seeks to provide insights into the mathematical properties, interpretation, and applicability of Makeham models in modeling age-dependent and age-independent mortality risks. Additionally, the paper aims to demonstrate formally that competing-risk models can be represented as mixture models, thereby facilitating a deeper understanding of risk-specific mortality dynamics.

[1] Interdisciplinary Centre on Population Dynamics, University of Southern Denmark, Odense, Denmark. Email: silca@sam.sdu.dk.
[2] Interdisciplinary Centre on Population Dynamics, University of Southern Denmark, Odense, Denmark. Email: trim@sam.sdu.dk.

CONTRIBUTION

Expressing competing-risk models as mixtures aids identifying the overall and age-specific share of deaths according to each of the competing risks. In particular, Makeham mortality models, when represented as mixtures, provide, first, a semantically and computationally convenient platform to disentangle age-dependent from age-independent mortality, and second, a straightforward specification that can easily be extended to account for unobserved heterogeneity. By expressing Makeham models as a convex combination of probability distributions, we are able to estimate the age-profile of age-independent mortality, especially at the oldest ages, at which we intuitively assume that most deaths are age-dependent (senescent). We are also able to estimate the senescent mortality component, which is the one to focus on when studying the aging process and its characteristics.

1. Introduction

Death can occur due to various causes, each characterized by its own failure time. Individuals are exposed to a finite number of such causes, and the interplay of the latter determines the observed length of life.

Let there be a finite number of causes of death[3] labeled $c_1, c_2, \ldots, c_n$. Suppose a non-negative random variable T_k with a hazard function $h_k(x)$ captures the age at death from cause c_k, $k = 1, 2, \ldots, n$, when c_k is the only cause of death. Failure time T_k measures the 'absolute potency' of cause c_k. The length of life for an individual is captured then by the random variable $Y = \min(T_1, \ldots, T_n)$, whose hazard function, $\mu(x)$, reflects the competing n cause-specific risks of dying (see, for example, Chiang 1961; Berman 1963; Gail 1975).

Makeham models are widely used for studying mortality as they account for a simple and straightforward competing-risk framework. They are characterized by a hazard function of the type

$$\mu(x) = h(x) + c\,, \tag{1}$$

where $h(x)$ is another hazard function. In this family of models, the hazard function $h(x)$ represents the age-dependent risk of death, often associated with the aging process, while the Makeham term c characterizes the age-independent risk. This structure allows for a

[3] We will use the terms 'causes of death' and 'causes' interchangeably to describe the exhaustive set of pathways leading to death. An individual can die from one of n such competing causes, where n can be finite or infinite. By 'causes of death' and 'causes' we do not necessarily refer to diseases according to WHO's International Classification of Diseases.

clear differentiation between risks that vary with age and those that do not. When

$$h(x) = ae^{bx}\,, \tag{2}$$

(1) reduces to the Gompertz–Makeham (GM) model (Gompertz 1825; Makeham 1860), while for

$$h(x) = \frac{ae^{bx}}{1 + \frac{a\gamma}{b}(e^{bx} - 1)}\,, \tag{3}$$

(1) yields the gamma-Gompertz–Makeham (ΓGM) model (Vaupel, Manton, and Stallard 1979). Beard (1959) and Kannisto (1994) provide alternative logistic curves with different asymptotes. The hazard $h(x)$ can be extended by an additive component that reflects infant and childhood mortality, as in Siler's model (Siler 1979)

$$h(x) = a_1 e^{-b_1 x} + ae^{bx}\,. \tag{4}$$

The Gompertz–Makeham, gamma-Gompertz–Makeham, and Siler models reflect a competing risk framework: An individual dies either as a result of biological processes at early or late ages or due to some extrinsic risk c, whatever strikes first. In this framework, c is often termed 'premature', capturing risks independent of age, which typically account for sudden, unexpected deaths not directly linked to aging. On the other hand, $h(x)$ is associated with the 'senescent' risk, reflecting the natural decline in health due to aging. In the Siler model, the negative exponential in $h(x)$ captures the infant mortality risk.

The Makeham term in all models described by the general framework in Equation (1) accounts for a competing exponentially distributed risk. Previous studies, such as Gail (1975), Elandt-Johnson (1976), and Hakulinen and Rahiala (1977), discuss mathematical properties and practical implications of assuming independence among failure times, as well as the consequences when this assumption is violated. They do not study, though, the representation of competing-risk models as mixtures, in which every distribution characterizes deaths due to a specific risk out of n competing ones.

Here, we demonstrate that a Makeham (and any other competing-risk) model can be represented as a mixture, that is, as a convex combination of probability distributions corresponding to dying from each cause of failure in the presence of the risks from all others. The weights in this convex combination capture the probabilities of each cause striking first. Representing Makeham models as mixtures provides insight into the interactions of various risk factors. By distinguishing among these risks, mixture models facilitate targeted health interventions, improve predictive capabilities, and aid elucidating complex mortality patterns.

2. The relationship

Suppose in a population death can occur due to n causes $c_1, c_2, \ldots, c_n$, and each individual is characterized by a corresponding random vector $\mathbf{T} = (T_1, T_2, \ldots, T_n)'$ representing the hypothetical failure time from each cause in the absence of other causes. Then, a non-negative continuous random variable Y, capturing the actual length of life for an individual, can be represented as $Y = \min(T_1, T_2, \ldots, T_n)$, its hazard function can be expressed as

$$h_Y(x) = h(x,1) + h(x,2) + \cdots + h(x,n)\,, \tag{5}$$

and its probability density function (PDF) can be represented as

$$f(x) = \sum_{j=1}^{n} h(x,j) \exp\left\{ -\sum_{k=1}^{n} H(x,k) \right\}, \tag{6}$$

where the hazard $h(x,k)$ captures the risk of dying due to cause c_k, given that no other competing cause struck before x, and $H(x,k) = \int_0^x h(t,k)dt$ is the corresponding kth cumulative sub-hazard function, $k = 1, \ldots, n$. (6) implies that the PDF of Y can be expressed as a mixture

$$f(x) = \sum_{j=1}^{n} \pi_j g_j(x)\,, \tag{7}$$

where $g_j(x) = f(x \mid j)$ denotes the PDF of the lifespan for individuals that die due to c_j, $j = 1, 2, \ldots, n$ while being exposed to all n risks, and π_j is the probability that cause c_j strikes first among all competing causes.

3. Proof

Proof. Denote the hazard function of T_j, $j = 1, \ldots, n$, by

$$h_j(t) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \mathbb{P}(t < T_j \leq t + \Delta | T_j > t)\,,$$

and the corresponding survival function by $S_j(t)$. The survival function of $\mathbf{T}$ is given by

$$S_{\mathbf{T}}(t_1, \cdots, t_n) = \mathbb{P}\left(\bigcap_{j=1}^{n} [T_j > t_j]\right), \qquad t_j \in (0, \infty), j = 1, \ldots, n\,,$$

which satisfies $S_{\mathbf{T}}(0, \cdots, 0) = 1$ and $S_{\mathbf{T}}(\infty, \cdots, \infty) = 0$. The function $S(\cdot)$ is continuous from the right and monotonically non-increasing in each argument.

Note that we cannot observe the failure times $T_1, \ldots, T_n$ simultaneously. As a result, $S_{\mathbf{T}}(t_1, \cdots, t_n)$ cannot be observed, nor can its form be tested (Elandt-Johnson 1976). Instead, as highlighted by Moeschberger and David (1971), what we observe is the individual lifetime denoted by $Y = \min(T_1, \ldots, T_n)$. The corresponding survival function is given by

$$S_Y(t) = \mathbb{P}\,(Y > t) = S_{\mathbf{T}}(t, \cdots, t)\,, \tag{8}$$

where t is the observed failure time.

With the assumption that $\mathbb{P}(T_i = T_j) = 0$ for all $i \neq j$, $i, j = 1, \ldots, n$, we can define the random variable J as the index of the smallest T_i. Therefore, $J = j$ implies that cause c_j is responsible for the death, and $T_j < T_i$ for all $i \neq j$, $i = 1, \ldots, n$. Applying the law of total probability and the Bayes's theorem to Equation (8), we get

$$S_Y(t) = \sum_{j=1}^{n} \mathbb{P}\,(Y > t, J = j) = \sum_{j=1}^{n} \mathbb{P}\,(Y > t \,|\, J = j)\,\mathbb{P}\,(J = j)\,.$$

The function

$$S(t \,|\, j) = \mathbb{P}\,(Y > t \,|\, J = j) = \frac{\mathbb{P}(t < T_j < \min_{i \neq j} T_i)}{\mathbb{P}(T_j < \min_{i \neq j} T_i)}$$

represents the cause-specific survival function for cause c_j, while the quantity

$$\pi_j = \mathbb{P}\,(J = j) = \mathbb{P}(T_j < \min_{i \neq j} T_i)$$

is the probability that cause c_j strikes first among all competing causes. Thus,

$$S_Y(t) = \sum_{j=1}^{n} \pi_j S(t \mid j)\,. \tag{9}$$

Assuming $S_{\mathbf{T}}$ is differentiable, we can express the density of Y as

$$f_Y(t) = -\frac{d}{dt} S_Y(t) = \sum_{j=1}^{n} f(t,j) = \sum_{j=1}^{n} \pi_j f(t \mid j)\,, \tag{10}$$

where $f(t \mid j) = -\frac{d}{dt} S(t \mid j)$ represents the density of lifetimes due to competing cause c_j. Integrating $f(t,j)$ with respect to t, we can derive π_j. The hazard function of Y is given by

$$h_Y(t) = \frac{f_Y(t)}{S_Y(t)} = \sum_{j=1}^{n} h(t,j)\,, \tag{11}$$

where $h(t,j) = \frac{f(t,j)}{S_Y(t)}$. This function is sometimes called the jth sub-hazard rate and should not be confused with the function $h(t \mid j) = -\frac{d}{dt} \log S(t \mid j) = f(t \mid j)/S(t \mid j)$.

Equations (9)–(11) illustrate that, no matter if the failure times $T_1, \ldots, T_n$ are independent or not, the distribution of lifetimes can be expressed as a convex combination of n probability distributions, each representing the time to death for individuals affected by a given cause c_j, $j = 1, \ldots, n$. □

Note 1: The assumption on the identity of the forces of mortality

As previously mentioned, the function $S_{\mathbf{T}}(t_1, \cdots, t_n)$ is not directly observable, nor can its form be tested. In practical applications, an additional assumption on the identity of the forces of mortality is typically made:

$$h(t,j) = h_j(t), \quad j = 1, \ldots, n. \tag{12}$$

This assumption implies that $S_Y(t)$ can be expressed as a product of the individual survival functions $S_j(t)$:

$$S_Y(t) = S_{\mathbf{T}}(t, \cdots, t) = \prod_{j=1}^{n} S_j(t)\,.$$

This assumption suggests that the events $[T_j > t]$, $j = 1, \ldots, n$ are independent. Note that this does not necessarily imply independence of the random variables $T_1, \ldots, T_n$ (see, for example, Gail 1975).

Assumption (12) simplifies the mathematical representation of the model and the statistical estimation of its parameters. It also ensures the model's identifiability, which is essential for uniquely determining the model parameters from the available data (see, for example, Kalbfleisch and Prentice 2011).

As noted by Elandt-Johnson (1976), directly observing the 'time-to-death from a given cause of death' random variable is not feasible, and estimating its associated survival function requires additional assumptions. In addition, Crowder (1991) demonstrates that even when the marginal survival functions $S_j(t_j)$ of the failure time T_j are known, observations of Y and J do not determine their joint survival function $S_{\boldsymbol{T}}$. Therefore, assuming independence between events like $[T_j > t]$, $j = 1, \ldots, n$, or even between the random variables T_i and T_j, for $i \neq j$, $j = 1, \ldots, n$, though not necessarily reflecting the true interaction mechanism among different causes of death, provides a reasonable framework for distinguishing between extrinsic and senescent mortality.

Note 2: On the independence assumption for cause-specific failure times

When we assume that the random variables $T_j, j = 1, 2, \ldots, n$, are mutually independent, the survival function $S_{\boldsymbol{T}}(t_1, \cdots, t_n)$ simplifies to $\prod_{j=1}^{n} S_j(t_j)$, making the proof of the relationship rather straightforward. However, assuming independence of cause-specific failure times has at least two implications: First, the associated cause-specific risks of dying act independently of one another, and second, eliminating one cause of death does not affect the force of mortality of the other cause. However, the adequacy of this assumption varies depending on the specific cause of death and its complex relationship with other causes. A comprehensive model should acknowledge that (a) a simple constant term, such as the Makeham term, may not accurately capture the complex mechanisms underlying premature deaths; (b) lifetimes may not be independent, highlighting the need for a more sophisticated modeling approach; and (c) the dependence of lifetimes, reflecting behavioral, social, and environmental risk factors, may vary across different age groups.

According to Cox (1959), when considering two risks, data collected on variables Y and J will not provide evidence contradicting the assumption that the failure times associated with these risks are independent of each other. In other words, observations of Y and J will not suggest dependence between the failure times linked to these risks.

Even when circumventing the assumption of independence of lifetimes by using the method proposed by Gumbel and Mustafi (1967) for random variables T_1 and T_2 following extreme value distributions, the estimation of correlation parameters is complex (see Tawn 1988). Thus, applying such a model to aggregate (life-table) mortality data is challenging.

Note 3: The case of an infinite number of causes

Note that the relationship holds also when the number of competing causes, denoted as n, is infinitely large. However, increasing the number of causes typically leads to an increase in the number of model parameters, which can pose significant challenges in accurately estimating the model (see, for example, Bühlmann and Van De Geer 2011). In practical implementations, where n is very large, this expansion of parameter space can intensify issues such as the dominance of certain risks over others, which can complicate the understanding of individual risks' impacts on the overall mortality distribution and pose substantial challenges in interpreting and applying the model results effectively.

4. Related results

Makeham models are a special case of Equation (5) when $n = 2$, $h_1(x) = c$ and $h_2(x) = h(x)$, as in Equation (1). This hazard function is comprised of additive competing risks of two causes of death: One of them, denoted by c_1, is age-independent, while the other, denoted by c_2, is age-dependent. Each cause is described by its own failure time, captured by the not necessarily independent random variables T_1 and T_2, respectively. The probability density function (PDF) of the actual lifetime $Y = \min(T_1, T_2)$ can then be represented as

$$f(x) = c \exp\{-cx - H(x)\} + h(x) \exp\{-cx - H(x)\}, \tag{13}$$

where $H(x) = \int_0^x h(t)dt$ is the cumulative hazard function of T_2. From Equation (13), it is easy to represent a Makeham model of the type in Equation (1) as a mixture model:

$$f(x) = \pi g_1(x) + (1 - \pi) g_2(x), \tag{14}$$

where $g_j(x) = f(x \mid j)$ denotes the PDF of the lifespan for individuals who die from c_j, $j = 1, 2$, being exposed to both the age-independent and the age-dependent risk of dying.

Each term in Equation (14) has a straightforward interpretation. For example,

$$\pi = \int_0^\infty c \exp\{-cx - H(x)\}\, dx\,, \tag{15}$$

known as the mixing proportion (McLachlan, Lee, and Rathnayake 2019), represents the probability of dying from the age-independent (premature) risk in the presence of a competing age-dependent (senescent) risk of dying, – that is, π is the premature mortality prevalence. On the other hand, the functions

$$g_1(x) = \frac{c}{\pi} \exp\{-cx - H(x)\} \quad \text{and} \quad g_2(x) = \frac{h(x)}{1-\pi} \exp\{-cx - H(x)\} \tag{16}$$

capture the PDFs of the distribution of deaths due to each of the two causes: In the presence of competing premature and senescent risks of dying, g_1 represents the PDF of premature deaths, while g_2 denotes the PDF of senescent deaths.

By representing the class of models in Equation (1) as a mixture of two probability distribution functions, we can determine the threshold age, x^*, that separates the age interval with prevailing premature deaths from the one with predominant senescent mortality:

$$x^* = \max\{\max\{x : g_1(x) = g_2(x)\}, 0\}\,. \tag{17}$$

Note that Equation (18) can alternatively be derived by knowing Equation (1) only:

$$x^* = \max\{\max\{x : h(x) = c\}, 0\}\,. \tag{18}$$

In Equation (18), x^* can be viewed as a critical transition point after which the age-dependent (senescent) risk overtakes the age-independent (external) one as the predominant risk. In a mixture model setting like Equation (17), x^* marks a shift in the predominant subpopulation risk. For Makeham models, it indicates a change in the dominant risk of dying, from external to senescent. In addition, mixture models account for the interplay between these risks over different age ranges.

From Equation (18), we can express the proportion of non-aging-related deaths at age x as

$$p(x) = c/(h(x) + c)\,. \tag{19}$$

$p(x)$ reflects the age-specific prevalence of premature mortality. Table 1 presents closed-form expressions for x^* and $p(x)$ for four popular Makeham mortality models.

Table 1: Expressions for the threshold age x^*, separating the age intervals with prevailing premature and senescent mortality, as well as the proportion of non-aging-related deaths at age x, $p(x)$, in Gompertz–Makeham, gamma-Gompertz–Makeham, Beard–Makeham, Kannisto–Makeham, and Siler model settings

Model	$h(x)$	x^*	$p(x)$
Gompertz–Makeham	ae^{bx}	$\frac{1}{b}\ln\frac{c}{a}$	$\frac{c}{ae^{bx}+c}$
Γ-Gompertz–Makeham	$\frac{ae^{bx}}{1+\gamma\frac{a}{b}\left(e^{bx}-1\right)}$	$\frac{1}{b}\ln\frac{c(b-a\gamma)}{a(b-c\gamma)}$	$\frac{c}{\frac{ae^{bx}}{1+\gamma\frac{a}{b}\left(e^{bx}-1\right)}+c}$
Beard–Makeham	$\frac{ae^{bx}}{1+kae^{bx}}$	$\frac{1}{b}\ln\frac{c}{a(1-kc)}$	$\frac{c}{\frac{ae^{bx}}{1+kae^{bx}}+c}$
Kannisto–Makeham	$\frac{ae^{bx}}{1+ae^{bx}}$	$\frac{1}{b}\ln\frac{c}{a(1-c)}$	$\frac{c}{\frac{ae^{bx}}{1+ae^{bx}}+c}$
Siler	$a_1e^{-b_1x}+a_2e^{b_2x}$	no closed-form	$\frac{c}{a_1e^{-b_1x}+a_2e^{b_2x}+c}$

Note that assuming a constant age-independent risk of dying does not imply a constant force of mortality for the distribution of premature deaths (in the presence of a competing senescent risk). The force of mortality for this subpopulation is given by

$$h(x\,|\,1)=\frac{\exp\{-cx-H(x)\}}{\int_x^\infty\exp\{-cy-H(x)\}\,dy}\,. \tag{20}$$

In many Makeham mortality models, we can derive closed-form expressions for different characteristics of the distributions of deaths in the mixture. For example, the remaining life expectancy at age x is given by

$$e_x=\frac{1}{S(x)}\int_0^\infty S(x+t)\,dt\,, \tag{21}$$

where $S(\cdot)$ is the survival function. Substituting $S(x)=S(x\,|\,1)=\int_x^\infty g_1(y)\,dy$ and $S(x)=S(x\,|\,2)=\int_x^\infty g_2(y)\,dy$ in Equation (21) yields e_x for the subpopulation struck by the premature and senescent competing risk, respectively. The modal age at death for these subpopulations corresponds to the maximum of $g_1(x)$ and $g_2(x)$, respectively. Note that for any non-decreasing $h(x)$, the modal age at death for the non-senescent subpopulation is zero. For the senescent subpopulation (with PDF g_2), the modal age at

death is the age where the overall force of mortality equals the relative derivative of $h(x)$ with respect to age:

$$h(x) + c = \frac{dh(x)/dx}{h(x)}\,. \tag{22}$$

For the Gompertz–Makeham model, if $b > c$, the senescent modal age at death is given by

$$M_{GM} = \frac{1}{b}\,\ln\frac{b-c}{a}\,, \tag{23}$$

while for the gamma-Gompertz–Makeham model, it is given by

$$M_{\Gamma GM} = \frac{1}{b}\,\ln\left(\frac{b}{a}\cdot\frac{b-a\gamma-c\left(1-\frac{a\gamma}{b}\right)}{b+a\gamma c}\right)\,. \tag{24}$$

When $c = 0$, the expressions for M_{GM} and $M_{\Gamma GM}$ reduce to the modal age at death for the Gompertz and the gamma-Gompertz model, respectively (for details, see Missov et al. 2015). The modal age at death can also be expressed for the Beard–Makeham model as

$$M_{BM} = \frac{1}{b}\,\ln\frac{b-c}{a(1+ck)}\,, \tag{25}$$

while in the Siler model setting, there is no closed-form expression. Note that, given the improvements in human mortality and the rise of longevity, we may observe a decreasing trend for c over time. This leads to a convergence of the overall modal age at death to the senescent one.

We can also derive closed-form expressions for π in a Gompertz–Makeham and a gamma-Gompertz–Makeham setting. From Equation (15), we can express $\pi = ce_0$, where e_0 is the life expectancy at birth of the Makeham model. Using closed-form expressions for e_0 (see, for example, Missov and Lenart 2013; Castellares et al. 2020; Castellares, Patricio, and Lemonte 2020), we can derive expressions for π in terms of special functions. For the Gompertz–Makeham model, we have that

$$\pi = \frac{c}{a}\left(\frac{a}{b}\right)^{\frac{c}{b}} e^{\frac{a}{b}}\,\Gamma\left(-\frac{c}{b},\frac{a}{b}\right)\,, \tag{26}$$

where $\Gamma(u,x) = \int_x^\infty t^{u-1}e^{-t}dx$, $x > 0$, $u \in \mathbb{R}$ is a complementary incomplete gamma function. For the gamma-Gompertz–Makeham model, the closed-form expression for π is given by

$$\pi = \frac{c\gamma}{b + c\gamma}\, {}_2F_1\left(\frac{1}{\gamma}, 1; \frac{1}{\gamma} + \frac{c}{b} + 1; 1 - \frac{a\gamma}{b}\right), \tag{27}$$

where ${}_2F_1(m,p;q;z) = \frac{\Gamma(q)}{\Gamma(p)\Gamma(q-p)} \int_0^1 u^{p-1}(1-u)^{q-p-1}(1-zu)^{-m}du$ is the Gaussian hypergeometric function.

We now focus explicitly on the Gompertz–Makeham model. The Makeham term has first been used to adjust Gompertz mortality estimates (Makeham 1860), resulting in the Gompertz–Makeham model. A non-negative continuous random variable Y has a Gompertz–Makeham distribution if its hazard function is given by

$$\mu(x) = ae^{bx} + c, \quad a, b > 0, \quad c \geq 0; \quad x \geq 0, \tag{28}$$

and survival function

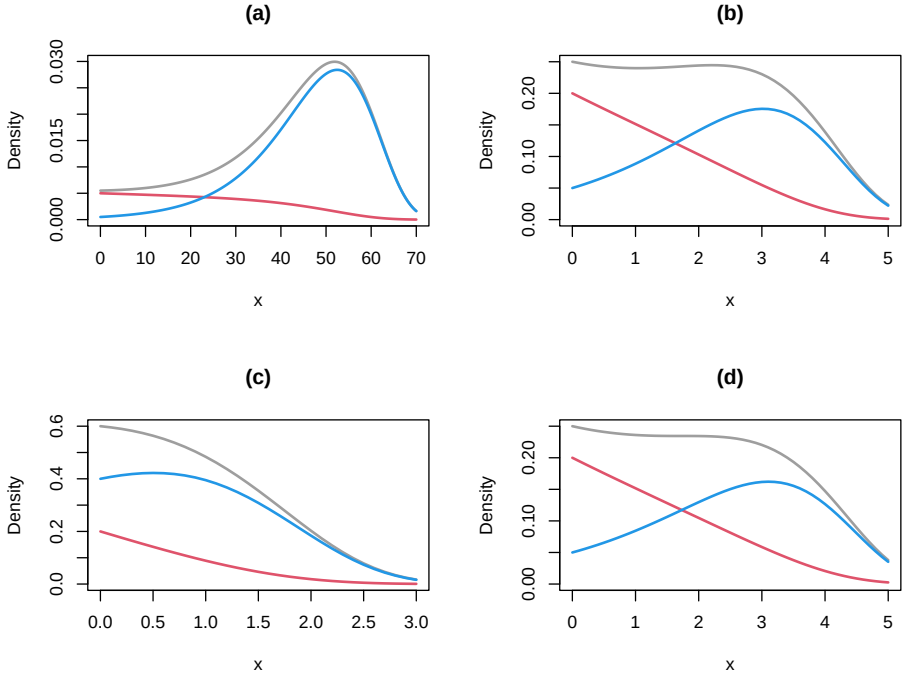
$$S(x) = \exp\left\{-\frac{a}{b}\left(e^{bx} - 1\right) - cx\right\}. \tag{29}$$

From Equations (28) and (29), it is straightforward to derive the Gompertz–Makeham PDF:

$$f(x) = c\exp\left\{-\frac{a}{b}\left(e^{bx} - 1\right) - cx\right\} + ae^{bx}\exp\left\{-\frac{a}{b}\left(e^{bx} - 1\right) - cx\right\}. \tag{30}$$

According to Castellares, Patricio, and Lemonte (2022), there are four different shapes for $f(x)$: (a) it may have a local maximum that is not located at the boundary; (b) It may have a local maximum both at the boundary and inside the parameter space; (c) it may have a global maximum at the boundary; and (d) it may have a global maximum at the boundary and an inflection point inside the parameter space (Norström 1997). The four shapes are displayed in Figure 1. While only panel (a) in Figure 1 corresponds to the pattern observed in human mortality, panels (b)–(d) were included to illustrate the diverse dynamics of mortality when decomposing senescence and premature mortality.

Figure 1: PDF of the Gompertz–Makeham distribution as a mixture



Notes: The gray line presents f, the red line indicates g_1, and the blue line represents g_2. For (a) we have $a = 0.0005$, $b = 0.1$, and $c = 0.005$; for (b) we have $a = 0.05$, $b = 0.85$, and $c = 0.2$; for (c) we have $a = 0.4$, $b = 0.8$, and $c = 0.2$; and for (d) we have $a = 0.05$, $b = 0.8$, and $c = 0.2$.

The distribution of a Gompertz–Makeham random variable Y can be perceived as a mixture – that is, its PDF in Equation (30) can be represented as a convex combination of two probability distributions

$$f(x) = \pi g_1(x) + (1 - \pi) g_2(x), \tag{31}$$

where

$$
\begin{aligned}
g_1(x) &= \frac{c}{\pi} \exp\left\{-\frac{a}{b}\left(e^{bx}-1\right) - cx\right\}, \\
g_2(x) &= \frac{ae^{-(c-b)x}}{1-\pi} \exp\left\{-\frac{a}{b}\left(e^{bx}-1\right)\right\}
\end{aligned}
\tag{32}
$$

and π is given by Equation (26).

Under the assumption of identity of the forces of mortality, the survival function takes the form

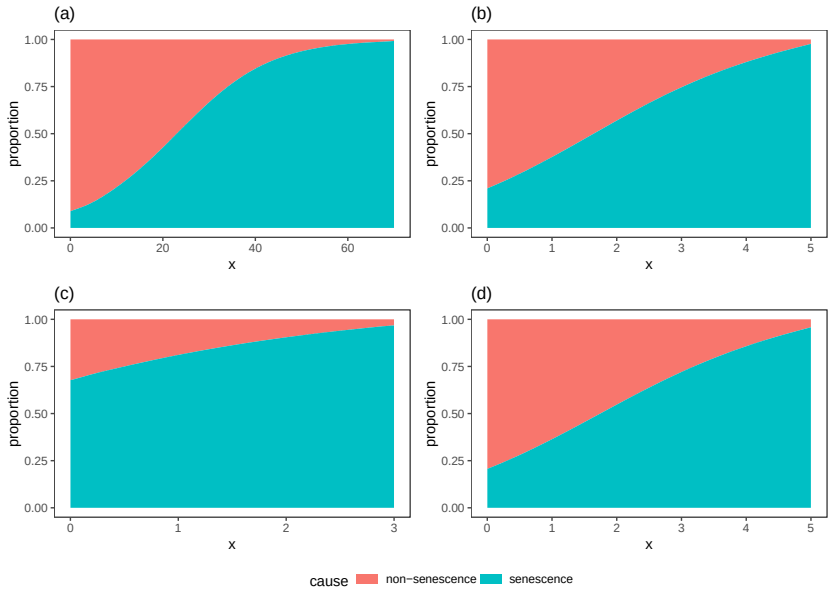
$$
S(x) = S_1(x)\, S_2(x) = \pi G_1(x) + (1-\pi) G_2(x), \tag{33}
$$

with $G_j(x) = S(x \mid j) = \int_x^\infty g_j(t)dt$ and $S_j(x) = \exp\left\{-\int_0^x h_j(t)dt\right\}$, for $j = 1, 2$.

Figure 1 shows the decomposition of the Gompertz–Makeham PDF into senescent (g_2 density) and non-senescent (g_1 density) components, from which we can see how non-senescent deaths shift the mode of the distribution (also known as 'modal age at death') to the left, sometimes even leading to a second peak of the distribution (panel (b) in Figure 1). Note that although the competing risks are Gompertz and Exponential, Equation (32) suggests that the Gompertz–Makeham density is not represented by a convex combination of a Gompertz and an Exponential distribution.

By representing the Gompertz–Makeham as a mixture of two distributions, we are able to quantify the overall proportion of non-senescent deaths, given by the quantity π, and also the proportion of non-senescent deaths at age x, equal to $p(x)$. For the Makeham–Gompertz model, the age-specific proportion of premature deaths can be expressed as $p(x) = c/\left(ae^{bx} + c\right)$. Figure 2 shows the proportion of senescent and non-senescent deaths by age in a Gompertz–Makeham model for different parameter combinations.

Figure 2: Proportion of senescent and non-senescent deaths by age in the Gompertz–Makeham setting



Notes: For (a) we have $a = 0.0005$, $b = 0.1$, and $c = 0.005$; for (b) we have $a = 0.05$, $b = 0.85$, and $c = 0.2$; for (c) we have $a = 0.4$, $b = 0.8$, and $c = 0.2$; and for (d) we have $a = 0.05$, $b = 0.8$, and $c = 0.2$.

The function $p(x)$ aids in estimating the threshold age x^* that separates the ages with predominant non-senescent deaths from the ages with predominant senescent deaths. For the Gompertz–Makeham model,

$$x^* = \begin{cases} \frac{1}{b} \ln \frac{c}{a} & c \geq a \\ 0 & c < a \end{cases} \tag{34}$$

when $x^* = 0$, $g_1(x)$ and $g_2(x)$ do not intersect (see, for example, panel (c) in Figure 2). In all other cases, x^* is their intersection point, (i.e., the point at which the proportion of senescent and non-senescent deaths is equal).

5. Applications

Expressing Makeham models in Equation (1) as a mixture can be advantageous for analyzing the components of mortality, (e.g., the ones capturing premature and senescent deaths). To illustrate these advantages, we estimate the Gompertz–Makeham model using raw death counts and exposures after age 20 from the Human Mortality Database (HMD 2023) for France, Italy, Japan, and Sweden, years 1947 to 2020, males and females separately.

Mazzuco, Suhrcke, and Zanotto (2021) highlight that the share of premature deaths is also defined by the shape, scale, and location of the senescent deaths distribution. This suggests a recommendation to maintain consistent senescent mortality parameters across countries while allowing for variability in premature mortality parameters. However, we will not keep the senescent mortality parameters constant for four main reasons: (1) The population in two different years is comprised of individuals from different cohorts with diverse life histories; (2) distinct genetic backgrounds, shaped by historical migrations and evolutionary processes, introduce variations in susceptibility to age-related diseases, potentially influencing senescent mortality; (3) the interplay between genetic factors and environmental determinants, including diet, lifestyle, pollution, and healthcare access, amplifies differences in senescent mortality patterns among countries; and (4) disparities in healthcare systems and socioeconomic factors add complexity, influencing disease prevention, diagnosis, and treatment effectiveness among countries.

To estimate the parameters, we apply a Bayesian procedure (Gelman et al. 2013) and assume a Poisson distribution for the death counts D_{ijk}, where the multi-index ijk represents age i in year j for country k (see, for example, Brillinger 1986). The Bayesian estimates are obtained by the mode of the posterior distributions, also known as the maximum a posteriori probability (MAP) estimate (Patricio and Missov 2023).

The prior (and hyper-prior) distributions are defined as

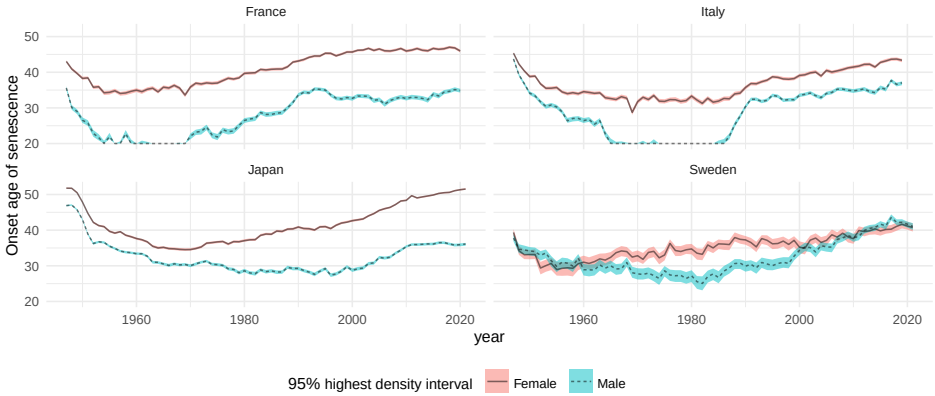
$$
\begin{aligned}
a_{ijk}|\alpha_1,\beta_1 &\sim InverseGamma(\alpha_1,\beta_1);\\
b_{ijk}|\alpha_2,\beta_2 &\sim InverseGamma(\alpha_2,\beta_2);\\
c_{ijk}|\alpha_3,\beta_3 &\sim InverseGamma(\alpha_3,\beta_3);\\
\alpha_1 &\sim Gamma(2,2);\\
\alpha_2 &\sim Gamma(2,2);\\
\alpha_3 &\sim Gamma(2,2);\\
\beta_1 &\sim Gamma(1,1);\\
\beta_2 &\sim Gamma(1,1);\\
\beta_3 &\sim Gamma(1,1),
\end{aligned}
$$

where a, b, and c are the Gompertz–Makeham model parameters. Since the parameters are strictly positive, we choose an inverse-gamma prior distribution. The latter is characterized by a heavy tail and keeps probability further from zero than the Gamma distribution.

By specifying the prior distributions, we employed the `NUTS` algorithm, a variant of Hamiltonian Monte Carlo known as the No-U-Turn Sampler (Hoffman and Gelman 2014; Betancourt 2017), via the `Rstan` R-package (Stan Development Team 2016) to sample from the posterior distributions. Four chains were executed, each consisting of 6,000 iterations (4,000 warm-ups and 2,000 sampling). We assessed convergence using the R-hat diagnostic and found that none of the R-hat values exceeded 1.05 (see, for instance, Vehtari et al. 2021). We opted not to thin the chains as thinning is typically unnecessary with the efficient Hamiltonian Monte Carlo method employed by Stan. The point estimates provided are MAP estimates, and the intervals displayed represent 95% highest density intervals for the posterior distributions.

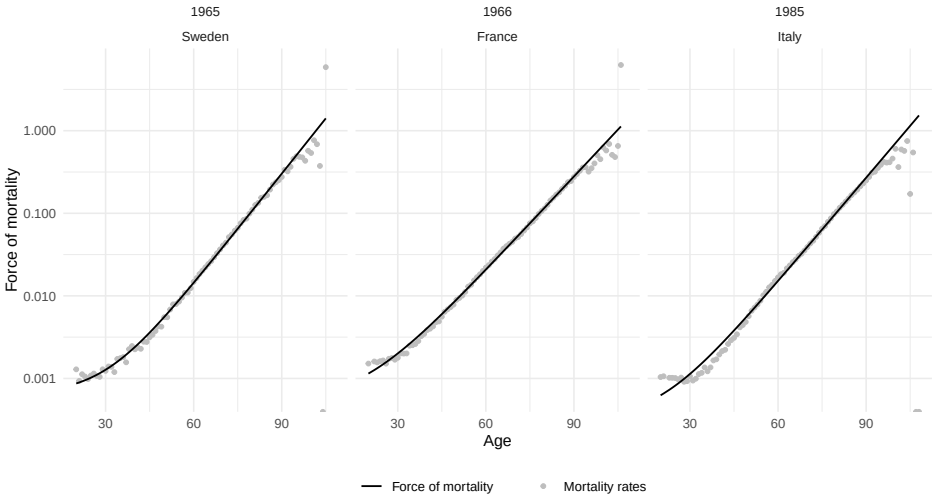
Figure 3 presents the threshold age of senescent mortality within the Gompertz–Makeham framework. In France, post-2000, the threshold age remains relatively stable, fluctuating between approximately 30 and 35 years for males and around 45 years for females. Almost the same holds for Japanese males after 2010. For other populations, notably after the 1990s, the threshold age exhibits a consistent linear increase. This trend suggests that in recent years, senescent mortality prevalence has been postponed at an approximate rate of three months per year for both Italian males and females, French males, and Japanese females. Swedish females experience a delay of around two months per year, while Swedish males see a more significant postponement of approximately four months per year. Japanese males and French females show a slower pace of postponement, at about half a month per year.

Figure 3: Threshold age between premature and senescence mortality estimated through the Gompertz–Makeham model



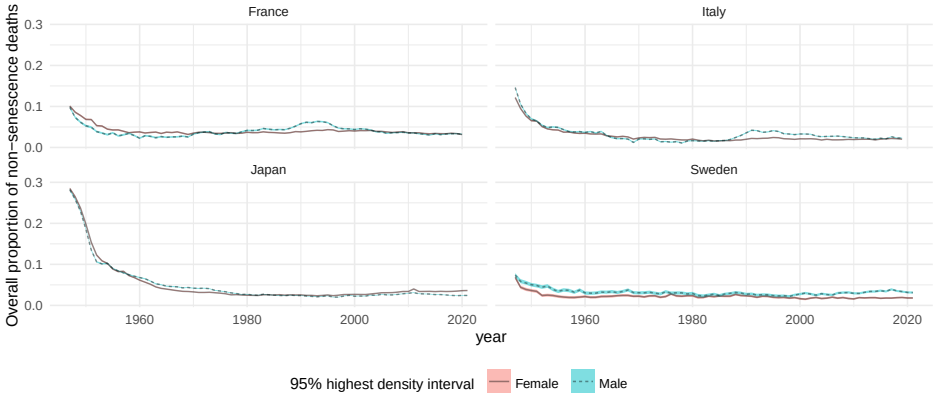
In Figure 3, we observe a consistently low threshold age for males in France (from 1960 to 1998) and Italy (from 1965 to 1985). During these periods, the data show a complexity that the Gompertz–Makeham model, shown in Figure 4, fails to capture. While the Gompertz–Makeham model is simplistic and fits well the log-linear increase in the risk of dying, it struggles to accommodate non-monotonic deviations from this pattern in the starting ages of analysis (Missov and Németh 2016). This limitation leads to an underestimation of the extrinsic risk parameter c. Consequently, while the Gompertz–Makeham model adequately represents the data pattern for Sweden, it notably fails to provide an accurate fit for France and Italy. This discrepancy is primarily attributed to the complex data pattern observed between ages 20 and 45, which deviates from both the log-linear increase in the risk of dying and the assumption of constant premature mortality risk. As a result, the estimated threshold age x^* is simply the starting age of analysis (see Figure 3).

Figure 4: Fitted versus observed mortality rates for males from Sweden in 1965, France in 1966, and Italy in 1985



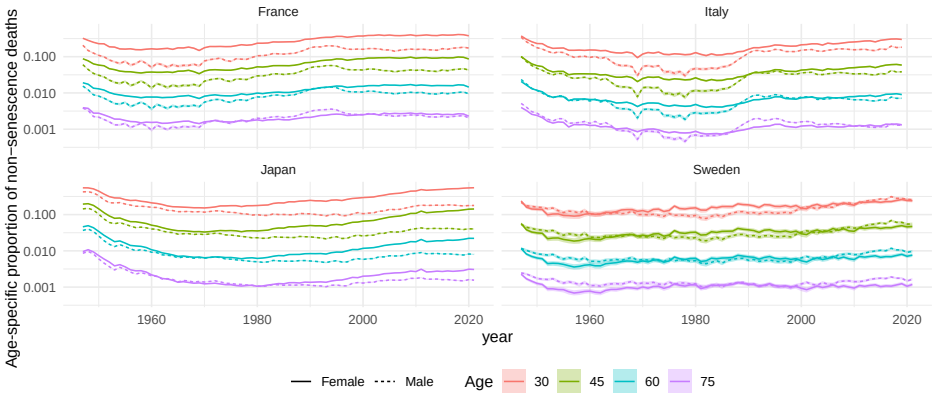
The prevalence of premature mortality after age 20 decreases over time (see Figure 5), faster until about 1970 and more moderately afterward, reaching an almost constant level in recent years. Sex-specific premature mortality seems to be balanced, with Sweden having a prevalence for males slightly higher than the one for females over the entire period. For French and Italian males, we observe a sudden increase from 1985 to 1990, leading to a persistent gap between male and female premature mortality prevalence until the late 1990s for France and about a decade later for Italy. This period coincides with the emergence of the global HIV/AIDS epidemic, impacting, in particular, France and Italy (see, for example Hamers et al. 1998). During this period, both countries experience a substantial rise in HIV/AIDS-related deaths. Meanwhile, although Sweden and Japan were also affected by the epidemic, the prevalence of HIV/AIDS was much lower compared to Italy and France (Gibney, DiClemente, and Vermund 2006). Although advancements in HIV/AIDS treatment emerged, access to these treatments varied. It was not until the later years of the 1990s that antiretroviral therapy became more widely accessible in France and Italy (Piot 2007; Ippolito et al. 2001).

Figure 5: Overall prevalence of premature mortality after age 20 estimated through the Gompertz–Makeham model



Despite the overall prevalence of premature mortality converging for some countries and diverging for others, when we look at the age-specific prevalence of premature mortality (Figure 6), we see a different picture. As expected, the prevalence of premature mortality decreases with age. However, for France, while the inter-sex difference in overall prevalence seems to converge to zero, we do not observe this for the age-specific prevalence at younger ages. The same trend is seen in Italy and Japan. For Sweden, however, what we see is the opposite. The prevalence seems to be the same for males and females between ages 30 and 45 and slightly higher for males from ages 60 to 75.

Figure 6: Age-specific prevalence of premature mortality estimated through the Gompertz–Makeham model



Over time, the age-specific prevalence of premature mortality for the French population seems to be stable for both sexes after 2000. The same holds for Japanese males after 2010 and for Italian males after 1990. For Sweden, the trend in age-specific prevalence of premature mortality seems to be identical for both sexes. Other mortality measures, such as the premature and senescent distributions of deaths, their force of mortality at ages 40, 60, and 80 and their modal age at death are presented in the appendix.

6. Conclusion

In this paper, we present a formal relationship that represents competing risk models as a mixture of distributions, where each cause of death corresponds to a component distribution, and an individual's lifespan is a mixture of these component distributions. The weights in this mixture correspond to the probabilities of each cause of death being the first to occur among all competing causes. This representation allows for a clear understanding of how competing risks influence overall mortality and provides a framework for modeling complex mortality patterns.

The mixture-model specification aids in representing the distribution of deaths as a convex combination of distributions for risk-specific subpopulations. This facilitates calculating various mortality and longevity measures for each subpopulation, as well as assessing the overall and age-specific prevalence of each cause of death.

We focus on the special case of Makeham mortality models that incorporate an age-independent risk of dying. We represent these models as a mixture that reflects individual lifetimes in a competing-risk setting. The interpretation is straightforward: An individual dies either according to a baseline mortality mechanism or an exponential distribution.

In the case of a two-risk model, the Gompertz–Makeham, when components capture senescent and premature deaths, we can estimate the threshold age that marks the change of death prevalence from premature to senescent. We are also able to reconstruct the age-specific profiles of the premature and senescent mortality components. To illustrate these findings, we take advantage of a Bayesian approach to estimate the Gompertz–Makeham model for the French, Italian, Japanese, and Swedish populations from 1947 to 2020 after age 20. The results suggest a postponement of senescent mortality prevalence at a pace between two and four months per year for most of the populations included in our analysis. The overall prevalence of premature mortality differs across age groups and between sexes.

7. Acknowledgments

The research leading to this publication is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant agreement No. 884328 – Unequal Lifespans). SCP gratefully acknowledges the support from the AXA Research Fund through funding the "AXA Chair in Longevity Research." The research and publication of this paper were supported by the SCOR Foundation for Science through the funding for the "SCOR Chair in Mortality Research." The article is also funded by the European Union (ERC, Born Once – Die Once, Grant agreement ID 101043983). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

Beard, R. (1959). Notes on some mathematical mortality models. In: Wolstenholme, G. and O'Connor, M. (eds.). *The lifespan of animals*. Boston: Little, Brown: 802–811.

Berman, S.M. (1963). Note on extreme values, competing risks and semi-Markov processes. *The Annals of Mathematical Statistics* 34(3): 1104–1106. doi:0.1214/aoms/1177704036.

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. ArXiv preprint arXiv:1701.02434.

Brillinger, D.R. (1986). A biometrics invited paper with discussion: The natural variability of vital rates and associated statistics. *Biometrics* 42(4): 693–734. doi:10.2307/2530689.

Bühlmann, P. and Van De Geer, S. (2011). *Statistics for high-dimensional data: Methods, theory and applications*. Berlin: Springer Science & Business Media. doi:10.1007/978-3-642-20192-9.

Castellares, F., Patricio, S., and Lemonte, A.J. (2022). On the Gompertz–Makeham law: A useful mortality model to deal with human mortality. *Brazilian Journal of Probability and Statistics* 36(3): 613–639. doi:10.1214/22-BJPS545.

Castellares, F., Patricio, S.C., and Lemonte, A.J. (2020). On gamma-Gompertz life expectancy. *Statistics and Probability Letters* 165: 108832. doi:10.1016/j.spl.2020.108832.

Castellares, F., Patricio, S.C., Lemonte, A.J., and Queiroz, B.L. (2020). On closed-form expressions to Gompertz–Makeham life expectancy. *Theoretical Population Biology* 134: 53–60. doi:10.1016/j.tpb.2020.04.005.

Chiang, C.L. (1961). On the probability of death from specific causes in the presence of competing risks. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, CA: University of California Press: 169–80.

Cox, D.R. (1959). The analysis of exponentially distributed life-times with two types of failure. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 21(2): 411–421. doi:10.1111/j.2517-6161.1959.tb00349.x.

Crowder, M. (1991). On the identifiability crisis in competing risks analysis. *Scandinavian Journal of Statistics* 18: 223–233.

De Moivre, A. (1731). *Annuities upon lives, or, the valuation of annuities upon any number of lives, as also, of reversions: To which is added, an appendix concerning the expectations of life, and probabilities of survivorship*. London: Samuel Fuller.

Elandt-Johnson, R.C. (1976). Conditional failure time distributions under competing risk theory with dependent failure times and proportional hazard rates. *Scandinavian Actuarial Journal* 1976(1): 37–51. doi:10.1080/03461238.1976.10405934.

Gail, M. (1975). A review and critique of some models used in competing risk analysis. *Biometrics* 31(1): 209–222. doi:10.2307/2529721.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., and Rubin, D.B. (2013). *Bayesian data analysis*. CRC press. doi:10.1201/b16018.

Gibney, L., DiClemente, R.J., and Vermund, S.H. (2006). *Preventing HIV in developing countries: Biomedical and behavioral approaches*. Berlin: Springer Science & Business Media.

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. F.R.S. &c. *Philosophical Transactions of the Royal Society of London* (115): 513–583. doi:10.1098/rstl.1825.0026.

Gumbel, E.J. and Mustafi, C.K. (1967). Some analytical properties of bivariate extremal distributions. *Journal of the American Statistical Association* 62(318): 569–588. doi:10.1080/01621459.1967.10482930.

Hakulinen, T. and Rahiala, M. (1977). An example on the risk dependence and additivity of intensities in the theory of competing risks. *Biometrics* 33(3): 557–559. doi:10.2307/2529375.

Hamers, F.F., Downs, A.M., Infuso, A., and Brunet, J.B. (1998). Diversity of the HIV/AIDS epidemic in Europe. *AIDS* 12: S63–S70.

Heligman, L. and Pollard, J. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries* 107(1): 49–80. doi:10.1017/S0020268100040257.

HMD (2023). The human mortality database. http://www.mortality.org/.

Hoffman, M.D. and Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15(1): 1593–1623.

Ippolito, G., Galati, V., Serraino, D., and Girardi, E. (2001). The changing picture of the HIV/AIDS epidemic. *Annals of the New York Academy of Sciences* 946(1): 1–12. doi:10.1111/j.1749-6632.2001.tb03899.x.

Kalbfleisch, J. and Prentice, R. (2011). *The statistical analysis of failure time data*. Hoboken, NJ: John Wiley & Sons. doi:10.1002/9781118032985.

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

Makeham, W.M. (1860). On the law of mortality and the construction of annuity tables. *Journal of the Institute of Actuaries* 8(6): 301–310. doi:10.1017/S204616580000126X.

Mazzuco, S., Suhrcke, M., and Zanotto, L. (2021). How to measure premature mortality? A proposal combining 'relative' and 'absolute' approaches. *Population Health Metrics* 19(41): 1–14. doi:10.1186/s12963-021-00267-y.

McLachlan, G.J., Lee, S.X., and Rathnayake, S.I. (2019). Finite mixture models. *Annual Review of Statistics and Its Application* 6: 355–378. doi:10.1146/annurev-statistics-031017-100325.

Missov, T. and Lenart, A. (2013). Gompertz–Makeham life expectancies: Expressions and applications. *Theoretical Population Biology* 90: 29–35. doi:10.1016/j.tpb.2013.09.013.

Missov, T.I., Lenart, A., Nemeth, L., Canudas-Romo, V., and Vaupel, J.W. (2015). The Gompertz force of mortality in terms of the modal age at death. *Demographic Research* 32(36): 1031–1048. doi:10.4054/DemRes.2015.32.36.

Missov, T.I. and Németh, L. (2016). Sensitivity of model-based human mortality measures to exclusion of the Makeham or the frailty parameter. *Genus* 71(2–3): 113–135. doi:10.4402/genus-665.

Moeschberger, M. and David, H. (1971). Life tests under competing causes of failure and the theory of competing risks. *Biometrics* 27(4): 909–933. doi:10.2307/2528828.

Norström, F. (1997). The Gompertz–Makeham distribution. Umeå: Umeå University (Master's thesis in Mathematical Statistics).

Patricio, S.C. and Missov, T.I. (2023). Using a penalized likelihood to detect mortality deceleration. *PLoS ONE* 18(11): e0294428. doi:10.1371/journal.pone.0294428.

Piot, P. (2007). *The HIV pandemic: Local and global implications*. Oxford: Oxford University Press.

Siler, W. (1979). A competing-risk model for animal mortality. *Ecology* 60(4): 750–757. doi:10.2307/1936612.

Stan Development Team (2016). Rstan: The R interface to stan. 522. https://mc-stan.org/. R package version 2.26.24.

Tawn, J.A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika* 75(3): 397–415. doi:10.1093/biomet/75.3.397.

Vaupel, J.W., Manton, K.G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16: 439–454. doi:10.2307/2061224.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.C. (2021). Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Analysis* 16(2): 667–718. doi:10.1214/20-BA1221.

Appendix

Figure A-1: Estimated force of mortality at ages 40, 60, and 80 through the Gompertz–Makeham model for premature and senescent groups

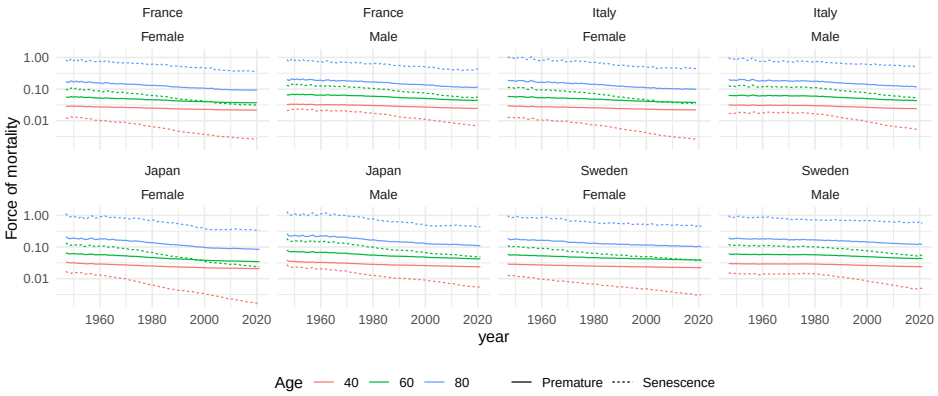


Figure A-2: Estimated senescence and overall modal age at death through the Gompertz–Makeham model

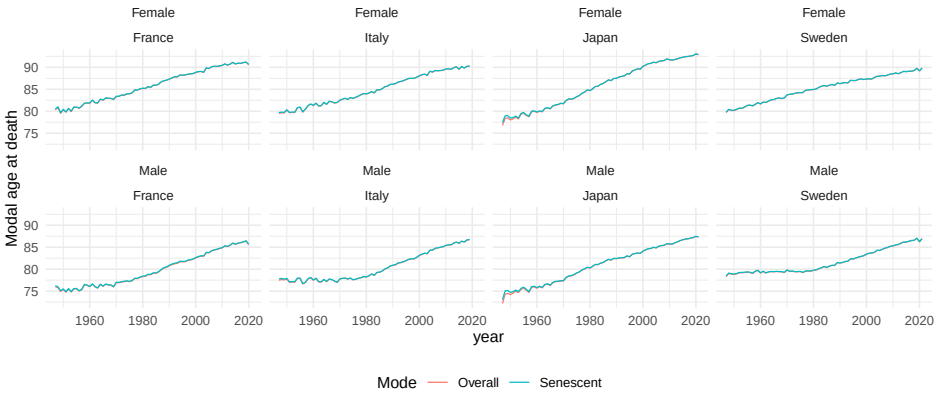


Figure A-3: Estimated non-senescent death distribution after age 20 through the Gompertz–Makeham model

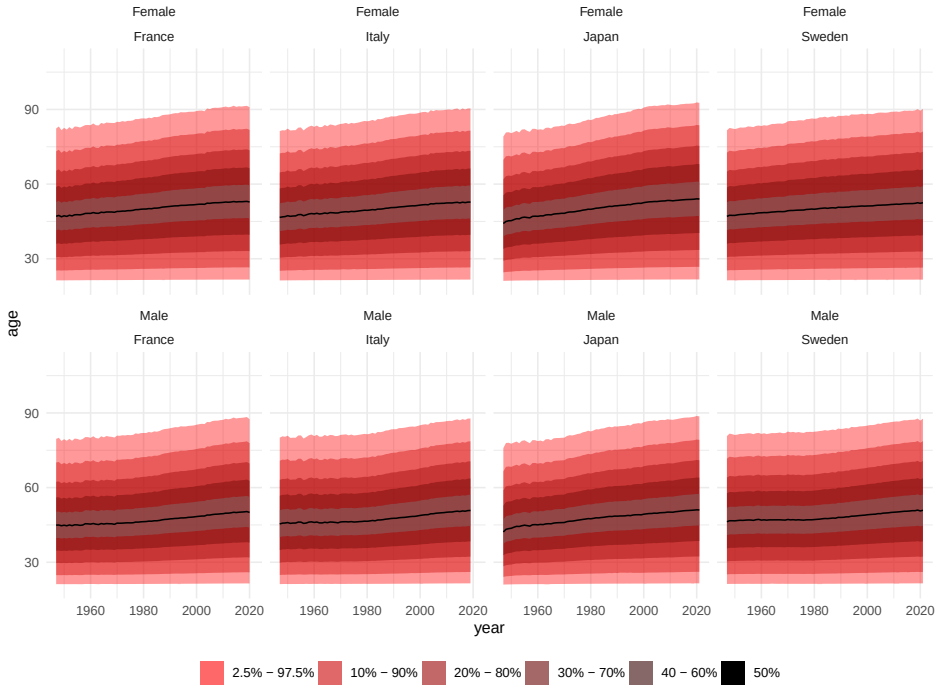


Figure A-4: Estimated non-senescent death distribution through the Gompertz–Makeham model