manuscript No.
(will be inserted by the editor)

The Age-Trajectory of Infant Mortality in the United States: Parametric Models and Generative Mechanisms

Jonas Schöley

Received: date / Accepted: date

Abstract Demographers are uncertain about the shape of infant mortality during the first year of life. While there is a consensus that the risk of death follows a Gompertz law over much of the adult age span, no such agreement exists about the parametric form of mortality at the very beginning of life with most literature on the topic suggesting either an exponential or some power-law expression. Based on individual level US birth and death register data I present evidence that the age-trajectory of infant mortality displays both power-law and exponential behavior and is better described by a product of those functions: an "exponentially-truncated power law". Across all infant populations under consideration the age-trajectory of mortality following birth is initially dominated by a power-law regime and over the course of infancy eventually approaches an exponential decline of less than a 1% reduction per additional day of age. The speed of the transition from a power-law to an exponential decline in the risk of death varies greatly by medical strata (5 minute APGAR score, birthweight and age of gestation at birth). The observed truncated power-law behavior may be the result of a shock-recovery or a mortality selection process.

Keywords Infant mortality · Infant life-table · APGAR score · Parametric survival analysis · Generalized linear model ·

During the writing of this article the author was a guest at the Max-Planck Institute for Demographic Research and funded by a grant from AXA Insurance.

Jonas Schöley
Interdisciplinary Centre on Population Dynamics, University of Southern Denmark
E-mail: jschoeley@health.sdu.dk

Background

Since Benjamin Gompertz published his eponymous law of mortality for the adult ages many suggestions have been made for a law of similar generality describing the age pattern of infant and/or childhood mortality. Maybe the earliest attempt was published by Oppermann (1870) who proposed[1] $\mu(x) = ax^{-1/2} + b + cx^{1/2}$ for mortality prior to age 20. With a power-law as first term this formula has the curious feature of predicting an infinite risk of death at the moment of birth, a property that reportedly was important to Oppermann though the reasons remain unclear (Steffensen 1930). One year later Thiele (1871), motivated by the search for a mortality law covering the whole human life-span assumed[2] $\mu(x) = ae^{-bx}$, the hazard of a negative Gompertz distribution, for the changing risk of death prior to maturity. *These two expressions, a power-law and an exponential, constitute the functional basis for most parametric models of infant/childhood mortality employed thereon.* Different power-law hazards have been used to describe the age pattern of infant mortality by Brillinger (1961), Choe (1981), de Beer and Janssen (2016) and Berrut et al. (2016); Gompertz-like exponential functions appeared as infant mortality terms in Siler (1979), Mode and Busby (1982), Rogers and Little (1994); and compositions of power- and exponential function have been suggested by Wittstein and Bumsted (1883) and Heligman and Pollard (1980).

The mortality models above can be interpreted as describing a not further specified risk to which an infant adapts over time, resulting in a continuous decline in mortality as the child grows (as explicitly stated by Siler 1979; Heligman and Pollard 1980).[3] Levitis (2011) called this the "acquired robustness hypothesis". A "competing-risks" explanation was proposed by Bourgeois-Pichat (1951) who hypothesized that the observed age pattern of mortality over the first year of life is the result of two separate processes: intrinsic mortality due to congenital disorders and extrinsic mortality due to accidents and maltreatment of the child. Noticing that the cumulative distribution of deaths during the post-neonatal period ($>$ 1 month of age) is closely matched by a linear function of $\log_{10}^3(\text{days since birth} + 1)$ while neonatal deaths follow a different age-trajectory Bourgeois-Pichat proposed that intrinsic and extrinsic mortality over age do not share the same functional form and that extrinsic mortality only starts to dominate over intrinsic mortality after the first month of life.

Some attempts have been made to express the parametric form of the infant hazard of death as a *frailty model* (Vaupel and Yashin 1983; Hougaard 1984; Vaupel and Yashin 1985). An important insight from these models is that in a population where individuals differ significantly in their risk of death the average mortality over age is determined not only by individual level age-effects but also by *mortality selection*, i.e. the changing composition of a cohort towards individuals with low frailty (Vaupel et al. 1979). It has been demonstrated that a *declining* infant mortality age-trajectory may result from *constant* individual level hazards of different magnitudes (Vaupel

[1] Throughout the paper $\mu(x)$ denotes the force of mortality at age x.

[2] ... after giving a full-page credit to Oppermann and only stopping short of apologizing for proposing a different expression...

[3] An exception is Oppermann (1870) who essentially proposed a competing risks model though the meaning Oppermann gave to the three components of his formula is unknown.

and Yashin 1983, 1985). Hougaard (1984) suggested that the high mortality right after birth and the fast subsequent decline may be the result of a large heterogeneity in individual frailties in the population of newborns.

In the presence of the ubiquitous semi-parametric Cox proportional hazards model (Cox 1972) and penalized smoothing spline models for count data (Eilers and Marx 1996; Camarda et al. 2016) it may seem like an exercise in nostalgia to consider parametric hazard expressions as a basis for the analysis of the infant mortality age pattern. Yet a parametric treatment of deaths during the first year of life comes with unique advantages:

1) Successes in combating infant mortality over the past century made infant death an increasingly rare event in large parts of the world (World Health Organization 2006; World Health Organization 2015). This challenges the statistician to develop a methodology suited to inference from rare events, especially if the object of study is infant mortality over age, on the regional level, by season, in nations with a small population, by socio-economic group, by cause-of-death, or by any combination of those criteria. If correctly specified, parametric survival models alleviate the statistical challenge of small event counts. By imposing a tight structure on the distribution of life-times one can gain insight from comparatively little data. Furthermore, a parametric specification of infant mortality expressed in terms of observable quantities (e.g. mortality at the day of birth, rate of post-neonatal mortality decline) facilitates the use of Bayesian methodology – a useful tool for working with sparse data – because informative prior distributions on the parameters are easily specified if the parameters are well understood.
2) Parametric models may allow insight into the mechanisms that determine the age-distribution of infant deaths. This is especially apparent with regards to the frailty model (Vaupel et al. 1979) which expresses the age-pattern of mortality on the population level as emerging from heterogeneous individual level hazards via selective mortality.
3) Information on the ages at death during infancy may only be available in broad age groups (e.g. birth to 1 day, 1 day to 1 week, 1 week to 1 month, months 1 to 12). If the parametric form of the ages at death during the first year of life is known this information can be used to graduate the grouped data. This is especially relevant for demographers who are interested a precise estimation of the "average age at death during infancy", a relevant statistic for the construction of life-tables.

The advantages above depend on the correct functional specification for the age distribution of infant deaths. Identifying this distribution requires a) data on the precise timing of each infant death in a birth cohort (ideally on a day-to-day basis); b) a cohort of infants large enough for a clear trend in the age-trajectory of infant mortality rates to dominate over random variation; and c) data on multiple birth-cohorts, and sub-populations so that it is possible to check the generality of the identified functional form. The individual level data on millions of births and infant deaths provided by the National Center for Health Statistics (2016) fits these requirements and for this article I confine the search for a "law" of infant mortality to the US population.

After a brief description of the data, the methods and the current day-to-day age pattern of infant mortality in the United States I will present evidence for both power-law *and* exponential behavior of the age-trajectory of infant death. Both types of models are then synthesized into a *truncated power-law* hazard expression, a generalization of both the power-law and negative Gompertz hazards, which fit is evaluated on life-tables for various cohorts and sub-populations of US infants. A discussion of the generative mechanisms which could give rise to a truncated power-law hazard concludes the paper.

Data

The "NCHS Cohort Linked Birth – Infant Death Data Files" (National Center for Health Statistics 2016) contain a complete census of births and infant deaths on the territory of the United States[4] and feature most fields present on the birth and infant-death certificates. The size and detail of the data allows for the calculation of day-to-day infant life-tables for various sub-populations.

I confine the analysis to the quinquennial birth cohorts 1995-1999 and 2005-2009 stratified by sex, prematurity, five minute APGAR score and social background of the mother. Those particular variables constitute major sources of heterogeneity in infant mortality along biological, medical and social dimensions and are among the most reliably reported fields.[5] In total I analyse data on 277,004 infant deaths out of 40,727,055 births over the birth cohorts 1995–1999 and 2005–2009.

Female and male infant life-tables are calculated for both birth cohorts, each life-table stratified by either the five minute APGAR score (an ordinal measure of an infants vitality shortly after delivery discretized into 3 levels), prematurity (defined via age of gestation upon delivery and discretized into 4 levels), mother's education (3 levels) or mothers ethnicity (4 levels). Those 56 life-tables thus cover heterogeneity in the age-trajectory of infant mortality among period, medical and social strata.

Methods

In order to determine the functional shape of infant mortality over the first year of life I fit the exponentially-truncated shifted-power family of hazards[6] to the day-to-day death counts and exposures for various populations of US infants. Let $\mu(x)$ be the

[4] Available data on the overseas territories has been excluded due to compatibility issues.

[5] It must be noted that the results of this paper can not be generalized to populations other than present day US infants without further study. While similar results for countries with similar overall levels of infant mortality (implying a similar level of development) are to be expected, it would be foolish to assume the same age pattern of infant mortality in pre-20th century populations or present populations suffering under a crisis: Historically the age-pattern of infant mortality was very much shaped by the interaction of seasonality effects, improper substitutes for breast-feeding, and deadly infectious diseases (Knodel and Kintner 1977; Huck 1995), factors which have lost relevance in the US over the course of the 20th century. Likewise an analysis of 21st century US infant mortality yields little insight into the characteristics of infant death in parts of the world suffering from humanitarian crises and violent conflicts.

[6] In the following referred to simply as "truncated power" hazard.

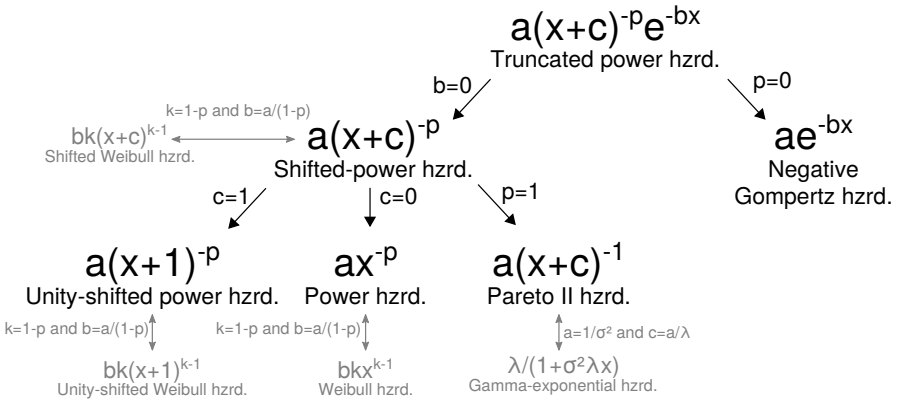

Fig. 1 The exponentially-truncated shifted-power family of hazards considered for the age-trajectory of infant mortality. The family contains the negative Gompertz, Pareto II, and Weibull hazards.

force of mortality (the hazard) at age x. The truncated power family is given by the expression

$$\mu_{\mathrm{TP}}(x) = a(x+c)^{-p}e^{-bx},$$

and contains many of the models proposed in the literature on infant mortality age-trajectories as special cases, such as the hazard functions of the negative Gompertz (Lomax 1954; Marshall and Olkin 2007, pp. 368) and the Pareto II distribution (Lomax 1954; Marshall and Olkin 2007, pp. 400), the exponential-Gamma frailty parametrization of the Pareto II distribution (Clayton and Cuzick 1985; Wienke 2011, p. 78), (shifted)-power hazards and their (shifted)-Weibull parametrizations (Lehman 1963; Marshall and Olkin 2007, pp. 321, see figure 1). Having all these models nested in a single family facilitates goodness-of-fit comparisons: A simpler model may be rejected in favor of a more complicated one on the grounds of significance test on the parameters and the whole family of models can be rejected if the full expression fails to accurately describe the data.

Prior to fitting the full truncated-hazard model to the data I will discuss the respective properties and parameter interpretation of the nested power- and negative Gompertz- hazards and demonstrate their fit (or lack thereof) to the observed death rates of the complete cohort of infants born in the US 2005 to 2009. This is to build evidence that it takes both a power- and an exponential component to adequately describe mortality over infancy, a subtlety that to my best knowledge has remained unmentioned by the literature so far.

In a second analysis I test the fit of the truncated power hazard by birth cohort, sex, APGAR score, gestation at birth, origin- and education of the mother. Confronting a mortality model with such a heterogeneous collection of life-tables acts as a validity test and allows for the quantification of heterogeneous mortality trajectories via comparison of model coefficients across population strata. Particular attention will

be paid to non-proportional variations of the hazard, and to the behavior of the hazard during the neonatal period (the first month of life) and during the post-neonatal period (month 1 to 12 of age).

All the models in this paper are fitted as generalized linear models (GLMs). Such an approach guarantees concave likelihood surfaces and consequently stable estimation of the model parameters while also capitalizing on the wide availability of software to fit and evaluate GLMs. By using suitable link functions and/or transformations of the age variable a range of parametric hazard functions can be linearized and thus fit to observed age specific death counts and exposures via a Poisson-GLM (Aitkin and Clayton 1980; Clayton 1983; Currie 2016).

The truncated power family of hazards (and therefore the nested negative Gompertz, Weibull and Pareto hazards, see figure 1) may be fit, after deciding on the value for c, as Poisson-GLMs with a log-link. This is possible due to the semi-linearity of $\mu_{\text{TP}}(x)$ on the log-scale:

$$\log \mu_{\text{TP}}(x) = \log(a) - p\log(x+c) - bx = \beta_0 + \beta_1 \log(x+c) + \beta_2 x,$$

with $a = \exp(\beta_0)$, $p = -\beta_1$ and $b = -\beta_2$. The non-linear c parameter is estimated by maximizing the *profile likelihood* (Murphy and Van der Vaart 2000) of c over a range of GLM fits.[7] Model inference is performed via the model deviances, Pearson-residuals, and a pseudo-R^2 criterion based on the fitted models deviance versus the deviance of the null (intercept-only) model. Confidence intervals around the parameter estimates of the truncated power hazard are calculated from 1000 repeated fits of the model on 1000 different parametric-bootstrap replicates of each life-table. Unlike the asymptotic standard errors retrieved from the GLM fit the bootstrap approach incorporates the uncertainty associated with the estimation of the non-linear c parameter.

The age-trajectory of US infant mortality

The age-trajectory of infant mortality as observed in the US birth cohort 2005-2009 is characterized by a peak right at birth with a rapid decline thereafter (see figure 2). While the mortality rate over the first hour following birth is around 0.02 (deaths per person-day of exposure), 24 hours later the hazard is at only 1.1 % of its initial value and then again drops by a factor of 10 over the next 29 days. Around one month after birth the hazard starts to decline with a near constant relative-rate, approaching an exponential behavior.

The pattern of a "super-exponential" decline in mortality during the neonatal period followed by an exponential tail over the remainder of the first year of life holds true for girls and boys, irrespective of the social background of their mothers, the APGAR score upon birth, the gestation at birth or birth cohort (see figures 4, 5, 6, 7).

[7] Let $\mathscr{L}(\beta, c)$ be the full likelihood of a GLM fit with coefficients β and age-offset c. The profile likelihood of c is given by $\mathscr{L}_p(c) = \max_\beta \mathscr{L}(\beta, c)$ and c is estimated as $\hat{c} = \arg\max_c [\max_\beta \mathscr{L}(\beta, c)]$. I found $\mathscr{L}_p(c)$ to be concave over c and the maximization was quick and stable.

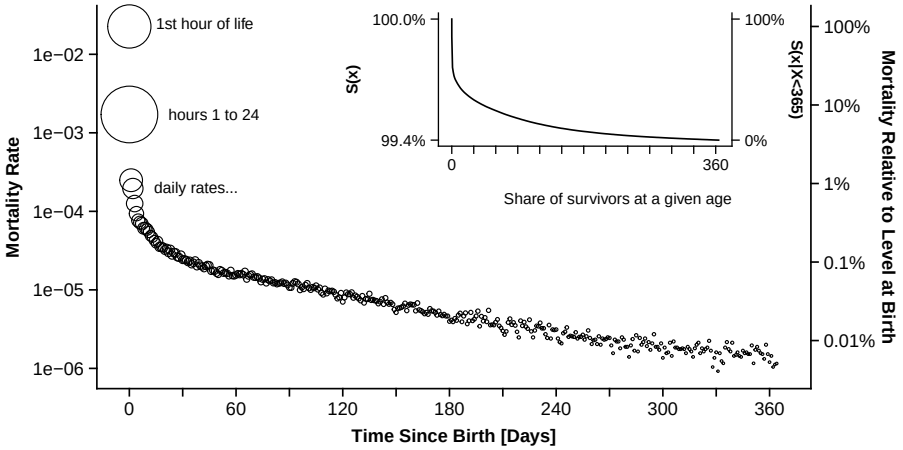

Fig. 2 Mortality rates over the first year of life for the US birth cohort 2005-2009. Mortality drops rapidly during the neonatal period after which it approches an exponential decline. Rates represent deaths per person-day of exposure. Circle area is proportional to the number of deaths at each interval. Raw data: NCHS Cohort Linked Birth - Infant Death Data Files.

Evidence for an exponentially-truncated power hazard

Negative Gompertz hazard

Thiele (1871), in an early attempt to model mortality across the whole human age-span, proposed to express the risk of death prior to maturity as an exponential function of age,

$$\mu_{\mathrm{GP}}(x) = ae^{-bx},$$

with $(x,a,b) \geq 0$. This is the hazard function of a Gompertz distribution with a negative b parameter and therefore mirrors the shape of the hazard commonly assumed for adult humans. The simplicity of an exponential term for the hazard facilitates analytic treatment and gives the parameters a direct interpretation as important quantities in the study of infant mortality, $\mu_{\mathrm{GP}}(0) = a$ being the hazard of death at the moment of birth and $-\frac{\mu'_{\mathrm{GP}}(x)}{\mu_{\mathrm{GP}}(x)} = b$ the instantaneous relative rate of mortality decline over age. The fact that the risk drops with a constant relative rate over age is the defining feature of the negative Gompertz hazard. Related quantities are independent of age as well: the time t it takes for the hazard to drop by a factor k is the solution to equation $\frac{\mu_{\mathrm{GP}}(x+t)}{\mu_{\mathrm{GP}}(x)} = \frac{1}{k}$ given by $t_{1/k} = \frac{\log(k)}{b}$ and for each unit increase in age the hazard changes by a factor $\frac{\mu_{\mathrm{GP}}(x+1)}{\mu_{\mathrm{GP}}(x)} = \exp(-b)$.

The negative Gompertz hazard can be expressed as a log-linear model with Poisson distributed age-specific death counts D_x as outcome and age-specific exposure times E_x as fixed offset[8]

[8] Assuming that the width of each age group is small enough as to not introduce substantial aggregation bias.

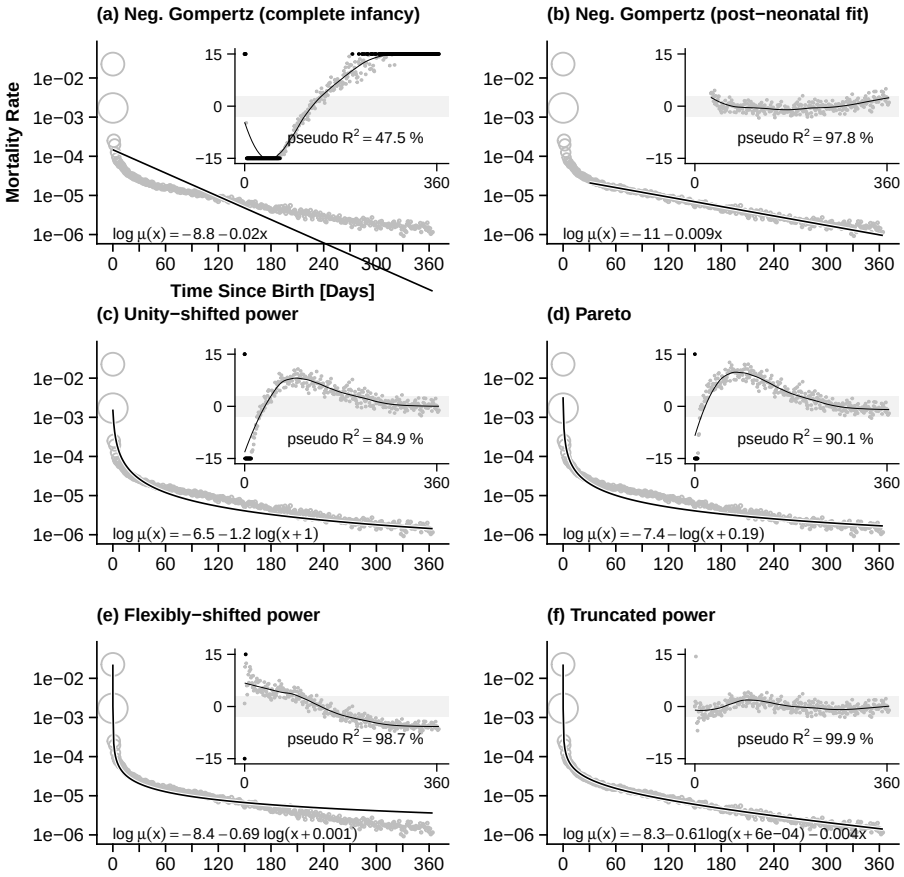

Fig. 3 Fitted versus observed mortality over the first year of life for the US birth cohort 2005-2009. The age-trajectory of infant mortality in the US exhibits both power-law and exponential behavior. While the negative Gompertz hazard with its exponential shape achieves the poorest of fits over the complete infancy period (a) it performs exceptionally well as a description for the post-neonatal mortality decline (b). Among the power-law expressions only the flexibly-shifted power hazard is capable of capturing the rapid decline in mortality over the first hours and days of life (e), still all three power-laws under consideration exhibit clear trends in their residuals over age (c, d, e). By combining a power-law and a negative Gompertz hazard in the truncated power model (f) we take into account the different behavior of the hazard in the neonatal- and post-neonatal periods and achieve a close fit over the entire first year of life. Rates represent deaths per person-day of exposure. Circle area is proportional to the number of deaths at each interval. R^2 is the percentage of deviance explained by the model. The Pearson-residuals mark the distance between prediction and observation in terms Poisson standard deviations around the prediction. Raw data: NCHS Cohort Linked Birth - Infant Death Data Files.

$$\log \mathrm{E}[D|x] = \beta_0 + \beta_1 x + \log E_x, \text{with } D \sim \text{Poisson}.$$

From this fit the parameters of μ_{GP} are recovered as $a = \exp(\beta_0)$ and $b = -\beta_1$.

The negative Gompertz model of infant mortality is especially popular in biology due to the influential paper by Siler (1979) who included it as part of his "Competing-

Table 1 Number of months over the five year period 2005-2009 a negative Gompertz fit achieved a lower deviance on the post-neonatal cohort life-tables compared to a given power-law fit. Data: US infants born in years 2005-2009. NCHS Cohort Linked Birth - Infant Death Data Files.

Gompertz vs.	Female	Male
Pareto II	44/60 73.3%	54/60 90%
Unity-shifted power	59/60 98.3%	60/60 100%
Flexibly-shifted power	23/60 38.3%	41/60 68.3%

Risk Model for Animal Mortality". In this model a Gompertz term with negative b parameter is used "to account for the hazard due to immaturity" (Siler 1979). For life-tables of human infants and children however, the negative Gompertz law has long been found to provide an insufficient fit (Thiele 1871, p. 326; Choe 1981; Gage and Dyke 1986).

Figure 3a clearly shows that the age-trajectory of infant mortality for the 2005-2009 US birth cohort deviates from a pure log-linear (i.e. exponential) form, ruling out the Gompertz hazard as a suitable model for the entire infant age range. *Yet*, a log-linear decline constitutes a near perfect description of the mortality trajectory over the post-neonatal period (figure 3b). The rate of death at day 30 after birth is around 2.03 deaths per 100,000 person-days of exposure with a subsequent decline of about 0.93% per additional day of age. Remarkably these two numbers explain 97.8% of the total deviance in the data during the post-neonatal period. Further investigating this result I fit negative Gompertz and power-law hazards to the post-neonatal life-tables for each of the 60 monthly US birth-cohorts from January 2005 to December 2009 separate by sex. In the majority of the cases the negative Gompertz hazard provided a closer fit to the observed post-neonatal mortality trajectory than either of the alternative two or three parameter power-law hazards (table 1).

Power-law hazard

Various power laws have been specified to characterize the age-specific hazard of infant death (Oppermann 1870; Brillinger 1961; Choe 1981; de Beer and Janssen 2016; Berrut et al. 2016) all of which are variations on the basic expression

$$\mu_{\mathrm{PW}}(x) = ax^{-p},$$

with $(a, p) \geq 0$ and $x > 0$. Just like a negative Gompertz hazard, a power-law hazard monotonically approaches 0 as $x \to \infty$, however, the relative rate of decline of a power hazard varies with age according to $-\frac{\mu'_{\mathrm{PW}}(x)}{\mu_{\mathrm{PW}}(x)} = \frac{p}{x}$, i.e. it is greatest right after birth and approaches 0 as age increases. It is this behavior that allows a power-law to initially decline faster than a Gompertz hazard, yet slower for larger x, rendering it potentially useful for the description of the day-to-day age pattern of infant mortality.

Similar to the Gompertz case the parameters of the power hazard can be interpreted in terms of mortality level and rate-of-change: The risk of death at age $x = 1$ is given by $\mu_{\mathrm{PW}}(1) = a$ and the proportional change in mortality for change in age

by proportion w is $\frac{\mu_{\mathrm{PW}}(wx)}{\mu_{\mathrm{PW}}(x)} = w^{-p}$. The proportional increase in age w needed for the hazard to drop by factor k is the solution to equation $\frac{\mu_{\mathrm{PW}}(wx)}{\mu_{\mathrm{PW}}(x)} = 1/k$ given by $t_{1/k} = k^{\frac{1}{p}}$. The power parameter p by itself represents the elasticity of the hazard, $\frac{\mathrm{d}\log \mu_{\mathrm{PW}}(x)}{\mathrm{d}\log x} = -p$, i.e. for every infinitesimal proportional increase in age the hazard drops by proportion p. A convenient property of power laws is that the elasticity is invariant to any re-scaling of age in the form $x' = wx$. In practice this means that the units used for the age column of the infant life-tables (e.g. hours, days, weeks) have no effect on the estimation of the exponent p.

Note that μ_{PW} can be re-parametrized into the hazard of the Weibull distribution $\mu_{\mathrm{UW}}(x) = bkx^{k-1}$ by substituting $p = 1-k$ and $a = bk$.

Berrut et al. (2016) fit μ_{PW} to death rates starting shortly after birth for a range of human and non-human populations and identify power-law behavior in most populations.[9] Specifically they identify a segmented power-law relationship between age and death rate for a cohort of Swiss infants. The first segment starts 1 hour after birth and lasts for 8 hours, while a different power coefficient is identified for the remainder of the first month of life.

Unity-shifted power hazard

As x approaches 0, hazard μ_{PW} approaches infinity. Due to this behavior, a pure power-law hazard can not be used to describe the age pattern of infant mortality starting from the moment of birth. Instead one can either choose to exclude the moment/hour/day of birth from the study period (as did Choe 1981; Berrut et al. 2016) or add a positive location parameter to μ_{PW}, i.e. $\mu_{\mathrm{PW}}(x+c)$. The location parameter can either be estimated from the data – allowing for additional model flexibility – or set to some constant. The latter option has the advantage that the resulting hazard can be written as a fully linear function of log-mortality. For mathematical convenience one may use a unity offset resulting in the *unity-shifted power* expression

$$\mu_{\mathrm{UP}}(x) = a(x+1)^{-p},$$

with $(x,a,p) \geq 0$. Due to the unity offset the risk of death at the moment of birth is $\mu_{\mathrm{UP}}(0) = a$. The elasticity of $\mu_{\mathrm{UP}}(x)$ approaches $-p$ as $x \to \infty$. In practice I found $-p$ to be a very close approximation to the true elasticity of $\mu_{\mathrm{UP}}(x)$ for at least the post-neonatal period thus one may safely interpret p as the approximate proportional drop in the hazard of infant death for an infinitesimal proportional change in age starting at day 30 after birth.

The unity-shifted power hazard can be fit as the log-linear Poisson regression

$$\log \mathrm{E}[D|x] = \beta_0 + \beta_1 \log(x+1) + \log E_x, \text{with } D \sim \text{Poisson},$$

with $a = \exp(\beta_0)$ and $p = -\beta_1$.

[9] Note that while occurrence-exposure rates are used in this article, Berrut et al. (2016) use the total number of births in the denominator of their death-rates.

Note that μ_{UP} can be translated into the unity-shifted Weibull hazard $\mu_{\mathrm{UW}}(x) = bk(x+1)^{k-1}$ by substituting $k = 1-p$ and $b = a/(1-p)$.

While the unity-shifted power-law fits the 2005-2009 US infant life-table better than the Gompertz hazard a look at the Pearson residuals in figure 3c still indicates a systematic misspecification. Due to its very nature the unity-shifted power-law can not capture the exponential portion of the hazard and it also fails to adequately describe the extremely fast drop in the risk of death over the first hours following birth.

Pareto hazard

In their model for the age pattern of human mortality de Beer and Janssen (2016) express the hazard of death during infancy and childhood as $\mu(x) = \frac{a}{c+x}$. This is the hazard function of a Pareto type II distribution (Lomax 1954; Marshall and Olkin 2007, pp. 400). Rewriting the hazard reveals a shifted power-law with exponent -1, scaling factor a and location offset c,

$$\mu_{\mathrm{PT}}(x) = a(x+c)^{-1},$$

where $(x,a) \geq 0$ and $c > 0$. While for the unity-shifted power-law the location offset is fixed and the power exponent varies the situation is reversed for the Pareto II hazard with a fixed power and a variable location offset. As $\mu_{\mathrm{PT}}(x)$ is finite for $x > -c$ one may interpret the c parameter as the duration μ_{PT} extends into the prenatal period. For such an interpretation to be justified the intrapartum death-rates (death during labor, i.e. at $x < 0$) have to follow the same functional form as the infant death rates and, given that $\mu(x+c)$ is declining with age, have to be higher than the mortality rates after birth. Of course the mortality at birth is $\mu_{\mathrm{PT}}(0) = a/c$ with a being the hazard level at $x = 1-c$. The elasticity of the Pareto hazard is $\frac{\mathrm{d}\log\mu_{\mathrm{PT}}(x)}{\mathrm{d}\log x} = -\frac{x}{x+c}$ which approaches -1 as $x \to \infty$. This implies a central characteristic of the Pareto hazard, namely that in the limit any proportional increase in age by factor w results in a proportional change of the hazard by $1/w$, e.g. doubling age halves the hazard. In practice the estimated values for c are small enough for this limiting behavior to set in shortly after birth.

The Pareto hazard can be fit as

$$\log \mathrm{E}[D|x] = \beta_0 - \log(x+c) + \log E_x, \text{ with } D \sim \text{Poisson},$$

an intercept-only model with additional offset $-\log(x+c)$ and $a = \exp(\beta_0)$. As described earlier c is estimated by maximizing its profile-likelihood.

Vaupel and Yashin (1983) propose to describe the infant/childhood component of human mortality by a Gamma-Exponential frailty model with population hazard function $\mu_{\mathrm{GE}}(x) = \lambda/(1+\sigma^2\lambda x)$, i.e. a continuous mixture distribution of constant baseline hazards (corresponding to an exponential baseline distribution of deaths) with Gamma-distributed rate parameter λ. As noted by Wienke (2011) this model is equivalent to μ_{PT} when $a = 1/\sigma^2$, i.e. the inverse of the variance parameter of the Gamma-Exponential frailty model, and $c = a/\lambda$. Therefore μ_{PT} may be interpreted as the population hazard resulting from mortality selection among individuals with constant hazards of Gamma-varying magnitudes.

The fit of the Pareto hazard to the 2005-2009 US infant life-tables is similar to that of the shifted-power-law: inadequate (figure 3d).

Flexibly-shifted power hazard

Adding a positive location offset c to μ_{PW} results in the flexibly-shifted power hazard

$$\mu_{\text{FP}}(x) = a(x+c)^{-p},$$

where $(x,a,b) \geq 0$ and $c > 0$. This hazard shape contains the Pareto II hazard and the unity-shifted power hazard as special cases. Note that μ_{FP} can be interpreted as the hazard function of the shifted Weibull distribution $\mu_{\text{SW}}(x) = bk(x+c)^{k-1}$ with $k = 1-p$ and $b = a/(1-p)$.

The elasticity of the flexibly-shifted power hazard is $\frac{\mathrm{d}\log\mu_{\text{FP}}(x)}{\mathrm{d}\log x} = -\frac{px}{x+c}$ which approaches $-p$ as $x \to \infty$. Due to small estimates for c this limiting behavior sets in shortly after birth, allowing an interpretation of the p parameter as the proportional drop in mortality for an infinitesimal proportional increase in age for all but the very first moments after birth. As for the Pareto II hazard parameter a is the hazard level at age $1-c$ and c may be interpreted as the time that μ_{FP} extends into the pre-natal period.

The flexibly-shifted power hazard can be fitted as a Poisson-GLM of the form

$$\log \mathrm{E}[D|x] = \beta_0 + \beta_1 \log(x+c) + \log E_x, \text{ with } D \sim \text{Poisson},$$

with the original parameters recovered as $a = \exp(\beta_0)$ and $p = -\beta_1$. Just like for the Pareto hazard c is estimated by maximizing its profile-likelihood.

Allowing for both a free power parameter p and a free location offset c greatly improves the fit to the US 2005-2009 infant life-tables compared to the more restricted unity-shifted power and Pareto hazards (figure 3e). The estimate for c is 0.0012 (i.e. a location shift by ≈ 1.7 minutes of age) and for p equals 0.69, corresponding to an approximate drop in mortality by $(1-2^{-0.69}) \times 100 = 38\%$ for every doubling of age. While fitting better than the hazards discussed above, the Pearson residuals in figure 3e still exhibit a clear curvature signifying a systematic lack of fit.

Exponentially-truncated power-law hazard

Multiplying the flexibly-shifted power-law hazard μ_{FP} with an exponential term results in the exponentially-truncated power expression

$$\mu_{\text{TP}}(x) = a(x+c)^{-p}e^{-bx},$$

with $(x,a,p,b) \geq 0$ and $c > 0$. This hazard contains all of the above models as special cases (see figure 1).

The power-component of the hazard is said to be "truncated" by the exponential component because with increasing age μ_{TP} approaches a constant relative rate of change, formally $\lim_{x\to\infty} \frac{\mu'_{\text{TP}}(x)}{\mu_{\text{TP}}(x)} = -b$.

Table 2 Percent reduction in residual deviance after truncating the power-law hazard with an exponential term. All results significant at $p < 0.001$ (via deviance ratio tests).

		1995-1999		2005-2009	
		Female	Male	Female	Male
APGAR	Low [5,9)	20.6	27.9	49.5	49.7
	Regular 9+	86.8	86.4	89.8	92.0
	Very low [0,5)	50.2	45.6	61.3	60.3
Education	College or university	67.3	66.0	76.6	76.8
	Elementary or less	38.6	41.9	30.5	40.9
	High school	76.9	74.7	79.9	77.3
Origin	Hispanic	55.4	63.6	60.2	69.4
	Non-Hispanic Black	56.0	60.1	65.1	63.9
	Non-Hispanic White	76.3	74.2	81.9	79.6
	Other	20.3	29.4	24.7	33.5
Prematurity	Extremely preterm ¡28w	67.4	67.1	74.1	75.7
	Moderate to late preterm [32,37)w	58.2	60.0	71.0	70.2
	Term 37w+	79.8	78.1	83.7	80.9
	Very preterm [28,32)w	54.7	56.1	60.9	63.0

The Poisson-GLM form of the truncated power hazard is

$$\log \mathrm{E}[D|x] = \beta_0 + \beta_1 \log(x+c) + \beta_2 x + \log E_x, \text{ with } D \sim \text{Poisson},$$

where $a = \exp(\beta_0)$, $p = -\beta_1$, $b = -\beta_2$ and the non-linear coefficient c is estimated via profile-likelihood maximization.

Adding the exponential term to μ_{FP} significantly ($p < 0.001$ via deviance-ratio test) improves the fit of the Poisson-GLM to the 2005-2009 US birth cohort and flattens the trend in the Pearson residuals over age (figure 3f). Infant mortality over age eventually follows an exponential trajectory – a result that, as will be demonstrated, holds true for different birth-cohorts, by sex, medical and social strata.

Evaluation of the truncated power hazard

The exponentially truncated power-law hazard achieves an excellent fit irrespective of cohort, sex, five minute APGAR score, gestational age at birth, origin or education of the mother. In all life-tables under consideration the percentage of deviance explained by the model ranges from 94.5% to 99.6%. For every single population the inclusion of an exponential term in addition to the shifted-power term significantly ($p < 0.001$ via deviance-ratio tests) improves the fit, reducing the residual deviance by more than 50% on 43 out of the 56 populations (table 2).

The shape of the hazard function over age varies most strongly by APGAR score (figure 4). The exponential behavior of the hazard is most pronounced for the population of infants born with a score of 9 or 10 (the majority of infants born). Within two weeks after delivery the initial mortality spike at birth transitions into a log-linear decline in mortality-risk of around 0.7% per day. Conversely, the hazard trajectory of

the life-table for infants with a "low" APGAR score (indicating health problems of upon delivery), while still exponential in the tail, features a more gradual transition into log-linear behavior. The greatest difference in the behavior of the hazard among the APGAR groups can be seen over the first week of life. While for APGAR group 0 to 5 the hazard drops by a factor of 1000 from the day of birth to an age of seven days it only drops by a factor of 9 or less for the other groups. Thus hazards vary by APGAR score in a highly *non-proportional* fashion as can also be inferred from the varying power (p), exponential-rate (b) and offset (c) parameters of the fitted models. Non-proportional behavior, albeit less pronounced, is also evident when comparing hazards by prematurity, with higher values for the power parameter $p = -\beta_1$ for lower ages of gestation at birth[10] (figure 6).

The relative rate of mortality decline during the post-neonatal period (as approximated by b) is higher for premature infants and infants with a low APGAR score compared to infants born at term or with a regular APGAR score. This implies that the relative difference in mortality between very frail and less frail infants diminishes over age to some degree.

Hazards are mostly proportional by ethnicity of the mother (figure 5) with the risk of death during the later stages of infancy declining exponentially with a rate of 0.3 to 0.5% per day of age and the power parameter ranging from -0.51 to -0.55 across the life-table strata. Notably during the post-neonatal period mortality over age declines slower for children of African-american mothers compared to infants of white mothers[11]; taken together with the observation that the intercept of the hazard curve is highest among infants of African-american mothers, this puts the group at a double disadvantage. Along education strata the hazards are proportional (figure 7).

Apart from a proportionally higher hazard for males compared to females there are no systematic differences in the age-trajectory of infant mortality between the sexes. Similarly the difference between the two birth cohorts is mostly proportional with the younger cohort having a lower intercept β_0 (figures 4–7).

Interpretation of the truncated power hazard

A shock-recovery process

The power-exponential product of the population hazard μ_{TP} can be interpreted as a *non-homogeneous split Poisson process*, where *shocks* to an infants health arrive with rate $\lambda(x)$ per unit person-time, each shock resulting in infant death with probability $p(x)$.[12]

Let $N(x)$ be the number of infants alive at age x and let $\mathrm{E}[M]$ be the expected value of a Poisson distributed random variable M with rate parameter $\int_x^{x+n} \lambda(x)N(x)$

[10] A result also observed by Berrut et al. (2016) for Swiss and Norwegian infants after fitting a simple power-law to age specific mortality rates.

[11] Significant at $p < 0.05$ for female and male infants of either birth cohort, see table 5 in the appendix and associated non-overlapping 95% confidence intervals around the b parameter for African-american and white mothers.

[12] Shock models in the context of human mortality have for example been studied by Strehler and Mildvan (1960), Finkelstein (2005), Cha and Finkelstein (2016).

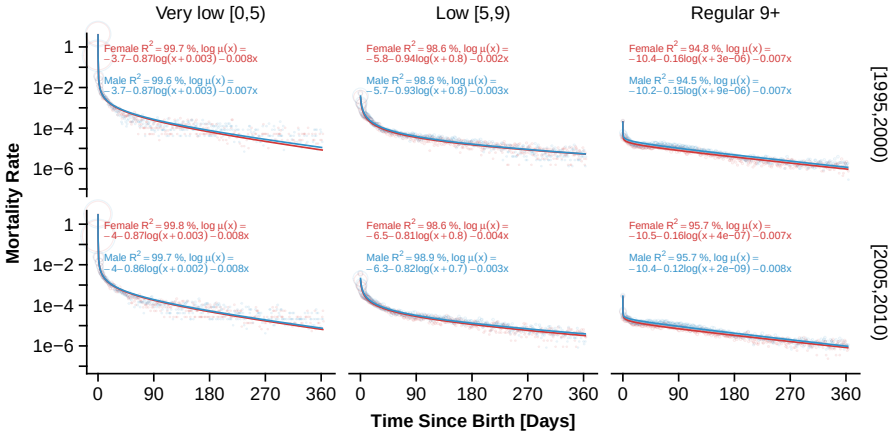

Fig. 4 Age specific hazard of death as predicted by the truncated power model contrasted with observed daily mortality rates by sex, birth-cohort, and 5-minute APGAR score. Rates represent deaths per person-day of exposure. Circle area is proportional to the number of deaths at each interval. R^2 is the percentage of deviance explained by the model. Raw data: NCHS Cohort Linked Birth - Infant Death Data Files.

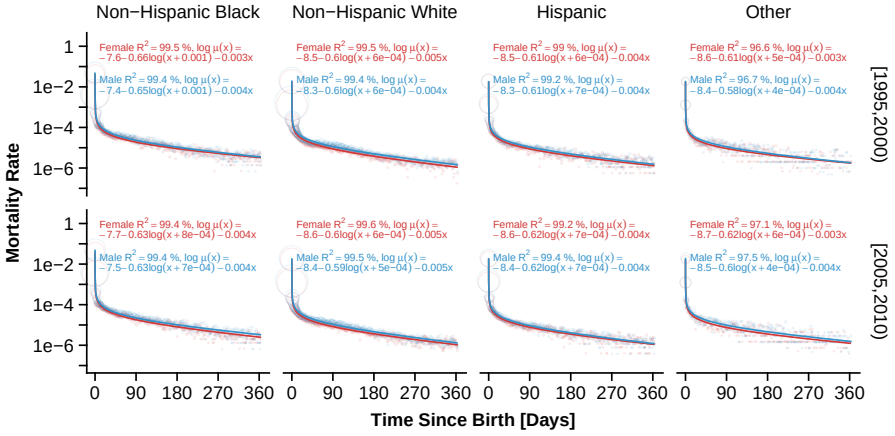

Fig. 5 Age specific hazard of death as predicted by the truncated power model contrasted with observed daily mortality rates by sex, birth-cohort, and ethnicity of the mother. Rates represent deaths per person-day of exposure. Circle area is proportional to the number of deaths at each interval. R^2 is the percentage of deviance explained by the model. Raw data: NCHS Cohort Linked Birth - Infant Death Data Files.

representing the total number of *health-shocks* the population of infants is expected to experience over age interval $[x, x+n)$. If each shock leads to death with probability $p(x)$ then the number of deaths D over age interval $[x, x+n)$ follows a Poisson distribution with expected value

$$\mathrm{E}[{}_nD_x] = \int_x^{x+n} \lambda(x)p(x)N(x)\,\mathrm{d}x,$$

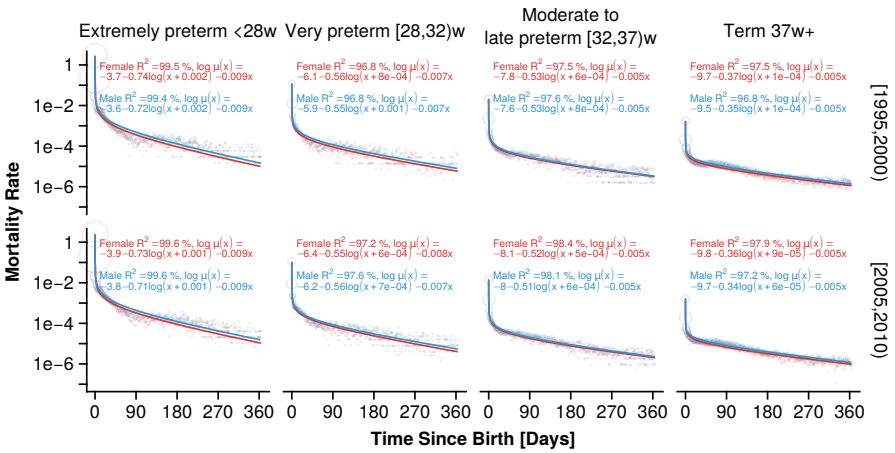

Fig. 6 Age specific hazard of death as predicted by the truncated power model contrasted with observed daily mortality rates by sex, birth-cohort, and prematuriy. Rates represent deaths per person-day of exposure. Circle area is proportional to the number of deaths at each interval. R^2 is the percentage of deviance explained by the model. Raw data: NCHS Cohort Linked Birth - Infant Death Data Files.

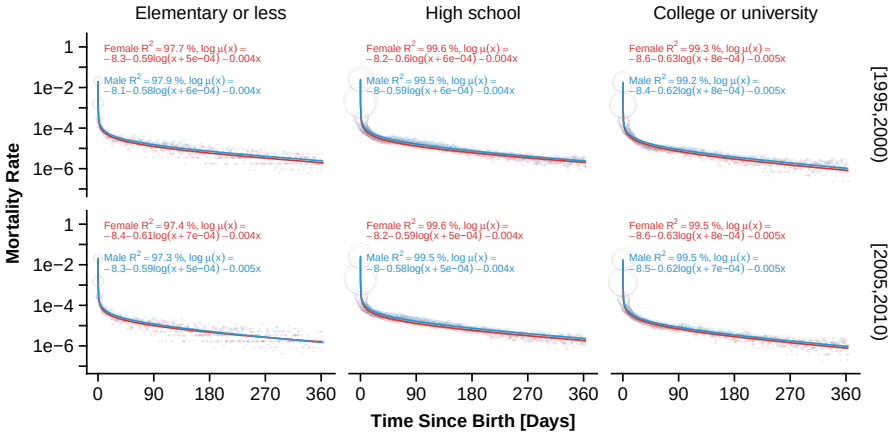

Fig. 7 Age specific hazard of death as predicted by the truncated power model contrasted with observed daily mortality rates by sex, birth-cohort, and education of mother. Rates represent deaths per person-day of exposure. Circle area is proportional to the number of deaths at each interval. R^2 is the percentage of deviance explained by the model. Raw data: NCHS Cohort Linked Birth - Infant Death Data Files.

see Prékopa (1958) for a proof. The hazard of death experienced by survivors N at time x is $\mu(x) = \lambda(x)p(x)$. If the rate of shocks $\lambda(x)$ varies over time according to a flexibly-shifted power hazard and if the probability of a shock leading to death $p(x)$ is exponentially declining the truncated power-law hazard is recovered. Note that neither the rate of shocks nor the probability of death following a shock are identified as this would require inferring the model $\mu(x) = a_1(x+c)^{-p} \times a_2 e^{-bx}$ from the fit $\mu(x) = a(x+c)^{-p}e^{-bx}$ – a problem with infinitely many solutions. Completely identified however are the power and the exponential rate parameters p and b which

can be interpreted as follows: For the extremely preterm female births of the US birth cohort 1995-1999 (figure 3f) the rate of shocks, $\lambda(x) \propto (x+0.002)^{-0.74}$, declined rapidly in the vicinity of birth by approximately $(1-2^{-0.74}) \times 100 = 40\,\%$ for every doubling of age whereas the probability of a shock leading to death, $p(x) \propto e^{-bx}$, declined by about $0.9\,\%$ per additional day of age.

Why use the power-law for the rate of shocks and the exponential term for the probability of a shock leading to death instead of the other way around – after all both options lead to the same expression for $\mu(x)$. I find it likely that the rapid neonatal mortality decline is predominantly a result of the waning stresses of birth – a sudden transition which may inflict a series of "shocks" to the infant – and not the result of a fast decline in the mortality risk associated with each shock. Furthermore, as $\lim_{x\to\infty} \frac{\mu'_{\mathrm{FP}}(x)}{\mu_{\mathrm{FP}}(x)} = 0$, the power-law approaches a constant hazard as age increases, allowing it to capture the age-independent rate of accidents which may dominate the rate of shocks later in infancy.

A very specific shock-recovery process is described by the Strehler-Mildvan model of senescent mortality. Strehler and Mildvan (1960) propose a model where individuals experience shocks to their health at a constant rate α. The magnitude of each shock, ranging from mild to severe, is a value drawn from an exponential random variable with rate β. The shocks are checked by the vitality of an individual which is a positive and declining function of age. When the magnitude of a shock at age x exceeds the individuals vitality at that age, death occurs. Strehler and Mildvan (1960) show that this process leads to a Gompertz distribution of life-times with corresponding hazard function $\mu(x) = \alpha e^{\beta x}$.

The truncated power hazard of infant mortality can be derived from such a Strehler-Mildvan process by changing two assumptions: 1) Where Strehler and Mildvan (1960) assume a decrease of vitality with time due to ageing one assumes an increase due to growth, and 2) instead of assuming shocks to arrive with a constant rate α one assumes that the risk of experiencing complications is highest at birth and falls over time (according to a shifted Weibull distribution distribution).

A mortality selection process

In the multiplicative frailty model (Vaupel et al. 1979) it is assumed that all individuals in a population share the same age specific "baseline hazard" of death $\mu_0(x)$ but on different "frailty" levels z which act multiplicatively on the baseline. Consequently the age specific hazard conditioned on frailty is given by

$$\mu(x|z) = z\mu_0(x).$$

Frailty is treated as a random variable Z with density $f(z|x)$ at age x. Integrating out the frailty yields the expression for the hazard observed at the population level, which is a mixture of the individual level hazards weighted by the age specific distribution of frailties

$$\overline{\mu}(x) = \int_0^\infty \mu(x|z) f(z|x)\,\mathrm{d}z,$$

which may be re-expressed as

$$\overline{\mu}(x) = \mu_0(x) \int_0^\infty z f(z|x)\,\mathrm{d}z = \mu_0(x)\mathrm{E}[Z|x],$$

showing that the population hazard is a product of the baseline hazard and the average frailty among the population at age x. By choosing $\mu_0(x) = ae^{-bx}$ and $\mathrm{E}[Z|x] = (x+c)^{-p}$ we express the truncated power-law hazard μ_{TP} as a multiplicative frailty model.

Frailty models can be interpreted in terms of *mortality selection*: Because individuals with a high frailty on average die earlier than those with lower values for z the average frailty among a cohort of individuals declines over time. Above we modeled this decline as an shifted power function of age and consequently p can be interpreted as the approximate elasticity of average frailty wrt. age, 0.69 for the US birth cohort 2005-1999, corresponding to an approximate drop in average frailty by 38 % for every doubling of age – a substantial mortality selection shortly after birth.

While differing in frailty, all individuals of a cohort are assumed to share the same baseline hazard. By modelling this *individual level* risk as an exponential function of age we assume that each additional unit of age decreases an infants risk of death by factor $\exp(-b)$, around 0.4 % per additional day of age for the 2005-1999 US birth cohort (3f). This relatively slow rate of decline may be attributed to *aquired robustness* due to the infants continued growth and development.

Of course one could, to the same effect, also choose $\mu_0(x) = a(x+c)^{-p}$ and $\mathrm{E}[Z|x] = e^{-bx}$. This model implies that the power-law behavior of the population hazard (mostly seen during the first month of life) arises from individual level processes while above we have assumed that the rapid mortality decline after birth is mostly due to changing average frailty, i.e. mortality selection. Different choices for the baseline hazard and the distribution of frailties can result in the same expression for the population hazard – an identifiability problem which is well known in the frailty-model literature (e.g. Hougaard 1995).

Discussion

Using highly detailed individual level data on the timing of death over the first year of life I found strong evidence for an exponentially truncated power-law behavior of the hazard of infant death in two recent US birth cohorts. The shift from a power-law regime to an exponential decline has not been noted in the literature before and invites speculation regarding the mechanisms that give rise to such a pattern. Observed hazards may be the result of mechanisms which have been discussed at length in the context of senescent risk of death: mortality selection due to heterogeneous frailties and a shock-recovery process.

The frailty hypothesis may be tested via a decomposition analysis. Vaupel and Zhang (2010) prove that, if the population hazard $\overline{\mu}(x)$ and the stratum specific hazards $\mu(x|z)$ are known for a cohort, then the rate of change over age of $\overline{\mu}(x)$ can be decomposed into a "direct change" component and change due to the process of mortality selection. Using this decomposition one could for example calculate how much

of the decline in the risk of death over the first week of life as shown in figure 2 is due the changing population composition by prematurity status resulting from the early death of extremely premature infants, i.e. mortality selection.

A test of the shock-recovery model is less straightforward as it requires both a clear definition of what constitutes a shock and data on the timing of such shocks to see if their rate of occurrence corresponds to either the exponential or the power-law component of the truncated power-law hazard.

Note that the truncated power hazard discussed in this paper does not permit an interpretation as a competing-risks model. Such a model was implied by Bourgeois-Pichat (1951) for the age-trajectory of infant mortality, which he partitioned into deaths due to congenital disorders and deaths due to accidents. An alternative competing-risks model featuring a power-law and an exponential term could take the form $\mu_{\mathrm{CR}}(x) = a_1(x+c)^{-p} + a_2 \exp(-bx)$, which may be re-written as $\mu_{\mathrm{CR}}(x) = \exp[\log a_1 - p\log(x+c)] + \exp[\log a_2 - bx]$. Sums-of-exponentials can be fit in the GLM framework by using a composite-link-function (Thompson and Baker 1981; Camarda et al. 2016). Assuming that deaths due to intrinsic causes follow a power-law behavior whereas extrinsic deaths feature the hazard of a negative-Gompertz distribution, this competing-risks formulation of the age-specific hazard of infant death over time permits a shape-based cause-of-death decomposition.

The results of this paper indicate that the proportional hazards assumption does not hold across APGAR score and prematurity strata. The association of those variables with the baseline hazard of death over the first year of life is highly non-proportional with much of the "effect" focussed on the first weeks of life. Indeed differences in the rate of post-neonatal mortality decline imply that the relative difference in mortality between very frail and less frail infants diminishes over age to some degree.

Insofar as the truncated power hazard generalizes to populations other than present day US it can become a valuable tool for working with low-quality data on the timing of infant deaths. Based on μ_{TP} one can calculate life-table a_0 from coarse data; design model infant-life-tables; smooth over data artifacts such as age-heaping; estimate seasonal effects on mortality separately from age effects; or include realistic infant mortality schedules into a simulation model.

Appendix: Parameter estimates and confidence intervals for the truncated power hazard

Table 3 The factor of mortality reduction over the first 7 days of life. Parameter estimates and 95% CIs calculated from the truncated-power hazard GLM fits for various US infant cohorts.

		1995-1999		2005-2009	
		Female	Male	Female	Male
APGAR	Low [5,9)	8 (6, 10)	8 (7, 10)	6 (5, 8)	7 (5, 8)
	Regular 9+	10 (7, 14)	8 (5, 10)	14 (11, 18)	13 (10, 16)
	Very low [0,5)	1006 (830, 1226)	965 (806, 1168)	1039 (880, 1253)	1007 (860, 1186)
Education	College or university	329 (291, 379)	302 (269, 341)	321 (283, 362)	315 (281, 355)
	Elementary or less	281 (207, 385)	235 (177, 312)	312 (224, 438)	291 (222, 393)
	High school	281 (254, 311)	257 (235, 282)	278 (250, 312)	273 (250, 299)
Origin	Hispanic	320 (263, 398)	277 (231, 330)	327 (278, 391)	309 (265, 359)
	Non-Hispanic Black	325 (281, 375)	317 (279, 361)	317 (274, 368)	339 (300, 388)
	Non-Hispanic White	303 (271, 339)	275 (250, 302)	301 (270, 337)	276 (250, 305)
	Other	333 (219, 499)	300 (213, 431)	347 (244, 496)	346 (258, 472)
Prematurity	Extremely preterm ¡28w	524 (456, 605)	449 (396, 510)	537 (474, 613)	477 (426, 531)
	Moderate to late preterm [32,37)w	148 (119, 186)	127 (103, 156)	143 (116, 176)	120 (101, 143)
	Term 37w+	58 (50, 68)	49 (43, 56)	59 (50, 69)	58 (50, 66)
	Very preterm [28,32)w	180 (134, 243)	145 (112, 187)	195 (150, 250)	177 (143, 221)

Table 4 Estimated a parameters of truncated-power hazard GLM fits for various US infant cohorts. Mean estimates and 95% confidence intervals are based on 1000 parametric bootstrap replications.

		1995-1999		2005-2009	
		Female	Male	Female	Male
APGAR	Low [5,9)	2.9e-03 (2.6e-03, 3.2e-03)	3.3e-03 (3.1e-03, 3.6e-03)	1.6e-03 (1.4e-03, 1.7e-03)	1.8e-03 (1.7e-03, 1.9e-03)
	Regular 9+	3.0e-05 (2.9e-05, 3.2e-05)	3.8e-05 (3.7e-05, 3.9e-05)	2.7e-05 (2.6e-05, 2.8e-05)	3.1e-05 (3.0e-05, 3.2e-05)
	Very low [0,5)	2.5e-02 (2.5e-02, 2.6e-02)	2.6e-02 (2.5e-02, 2.6e-02)	1.8e-02 (1.8e-02, 1.8e-02)	1.8e-02 (1.7e-02, 1.8e-02)
Education	College or university	1.9e-04 (1.8e-04, 1.9e-04)	2.2e-04 (2.2e-04, 2.3e-04)	1.8e-04 (1.7e-04, 1.8e-04)	2.1e-04 (2.0e-04, 2.1e-04)
	Elementary or less	2.5e-04 (2.4e-04, 2.6e-04)	2.9e-04 (2.8e-04, 3.0e-04)	2.2e-04 (2.1e-04, 2.3e-04)	2.6e-04 (2.5e-04, 2.7e-04)
	High school	2.7e-04 (2.7e-04, 2.8e-04)	3.4e-04 (3.4e-04, 3.4e-04)	2.6e-04 (2.6e-04, 2.7e-04)	3.3e-04 (3.2e-04, 3.3e-04)
Origin	Hispanic	2.0e-04 (1.9e-04, 2.0e-04)	2.4e-04 (2.3e-04, 2.4e-04)	1.8e-04 (1.8e-04, 1.9e-04)	2.2e-04 (2.2e-04, 2.3e-04)
	Non-Hispanic Black	4.8e-04 (4.7e-04, 4.9e-04)	5.8e-04 (5.7e-04, 5.9e-04)	4.5e-04 (4.4e-04, 4.6e-04)	5.4e-04 (5.3e-04, 5.5e-04)
	Non-Hispanic White	2.0e-04 (1.9e-04, 2.0e-04)	2.5e-04 (2.4e-04, 2.5e-04)	1.9e-04 (1.8e-04, 1.9e-04)	2.3e-04 (2.3e-04, 2.3e-04)
	Other	1.8e-04 (1.8e-04, 1.9e-04)	2.1e-04 (2.0e-04, 2.2e-04)	1.7e-04 (1.6e-04, 1.8e-04)	2.0e-04 (1.9e-04, 2.1e-04)
Prematurity	Extremely preterm ¡28w	2.4e-02 (2.3e-02, 2.4e-02)	2.8e-02 (2.8e-02, 2.8e-02)	2.0e-02 (1.9e-02, 2.0e-02)	2.3e-02 (2.3e-02, 2.3e-02)
	Moderate to late preterm [32,37)w	4.3e-04 (4.1e-04, 4.4e-04)	4.9e-04 (4.8e-04, 5.0e-04)	3.0e-04 (2.9e-04, 3.0e-04)	3.2e-04 (3.2e-04, 3.3e-04)
	Term 37w+	6.3e-05 (6.2e-05, 6.5e-05)	7.5e-05 (7.3e-05, 7.6e-05)	5.3e-05 (5.2e-05, 5.5e-05)	6.1e-05 (6.0e-05, 6.3e-05)
	Very preterm [28,32)w	2.2e-03 (2.1e-03, 2.3e-03)	2.7e-03 (2.6e-03, 2.8e-03)	1.7e-03 (1.6e-03, 1.8e-03)	2.0e-03 (1.9e-03, 2.1e-03)

Table 5 Estimated b parameters of truncated-power hazard GLM fits for various US infant cohorts. Mean estimates and 95% confidence intervals are based on 1000 parametric bootstrap replications.

		1995-1999		2005-2009	
		Female	Male	Female	Male
APGAR	Low [5,9)	2.2e-03 (1.7e-03, 2.7e-03)	2.6e-03 (2.1e-03, 3.0e-03)	3.9e-03 (3.4e-03, 4.4e-03)	3.5e-03 (3.1e-03, 3.9e-03)
	Regular 9+	7.0e-03 (6.7e-03, 7.2e-03)	7.2e-03 (6.9e-03, 7.4e-03)	7.1e-03 (6.8e-03, 7.3e-03)	7.7e-03 (7.5e-03, 7.9e-03)
	Very low [0,5)	8.0e-03 (7.2e-03, 8.8e-03)	7.3e-03 (6.7e-03, 8.0e-03)	7.7e-03 (7.1e-03, 8.4e-03)	7.6e-03 (7.0e-03, 8.2e-03)
Education	College or university	4.7e-03 (4.4e-03, 4.9e-03)	4.7e-03 (4.4e-03, 4.9e-03)	4.8e-03 (4.5e-03, 5.0e-03)	4.8e-03 (4.5e-03, 5.0e-03)
	Elementary or less	3.8e-03 (3.2e-03, 4.3e-03)	3.8e-03 (3.2e-03, 4.3e-03)	3.7e-03 (3.1e-03, 4.3e-03)	4.7e-03 (4.1e-03, 5.3e-03)
	High school	3.7e-03 (3.5e-03, 3.9e-03)	4.0e-03 (3.8e-03, 4.2e-03)	4.1e-03 (3.9e-03, 4.3e-03)	4.2e-03 (4.0e-03, 4.3e-03)
Origin	Hispanic	3.9e-03 (3.5e-03, 4.2e-03)	4.1e-03 (3.7e-03, 4.4e-03)	4.0e-03 (3.6e-03, 4.3e-03)	4.4e-03 (4.0e-03, 4.7e-03)
	Non-Hispanic Black	3.1e-03 (2.8e-03, 3.4e-03)	3.6e-03 (3.4e-03, 3.8e-03)	4.1e-03 (3.8e-03, 4.3e-03)	3.8e-03 (3.5e-03, 4.0e-03)
	Non-Hispanic White	4.5e-03 (4.3e-03, 4.7e-03)	4.5e-03 (4.3e-03, 4.7e-03)	4.5e-03 (4.3e-03, 4.7e-03)	4.7e-03 (4.5e-03, 4.9e-03)
	Other	3.0e-03 (2.4e-03, 3.7e-03)	3.7e-03 (3.1e-03, 4.3e-03)	3.4e-03 (2.8e-03, 4.1e-03)	3.7e-03 (3.2e-03, 4.3e-03)
Prematurity	Extremely preterm ¡28w	9.4e-03 (8.9e-03, 9.9e-03)	9.2e-03 (8.8e-03, 9.6e-03)	8.8e-03 (8.3e-03, 9.3e-03)	8.6e-03 (8.3e-03, 9.0e-03)
	Moderate to late preterm [32,37)w	5.0e-03 (4.6e-03, 5.4e-03)	5.1e-03 (4.8e-03, 5.5e-03)	5.3e-03 (4.9e-03, 5.7e-03)	5.4e-03 (5.1e-03, 5.7e-03)
	Term 37w+	5.1e-03 (4.9e-03, 5.3e-03)	5.3e-03 (5.1e-03, 5.5e-03)	5.2e-03 (5.0e-03, 5.4e-03)	5.3e-03 (5.1e-03, 5.5e-03)
	Very preterm [28,32)w	7.2e-03 (6.5e-03, 7.9e-03)	7.0e-03 (6.4e-03, 7.6e-03)	7.7e-03 (7.1e-03, 8.3e-03)	7.2e-03 (6.7e-03, 7.7e-03)

Table 6 Estimated c parameters of truncated-power hazard GLM fits for various US infant cohorts. Mean estimates and 95% confidence intervals are based on 1000 parametric bootstrap replications.

		1995-1999		2005-2009	
		Female	Male	Female	Male
APGAR	Low [5,9)	8.1e-01 (7.1e-01, 9.2e-01)	7.9e-01 (7.0e-01, 8.8e-01)	8.0e-01 (6.9e-01, 9.2e-01)	7.4e-01 (6.5e-01, 8.3e-01)
	Regular 9+	5.6e-06 (4.0e-07, 2.4e-05)	1.5e-05 (1.1e-06, 6.9e-05)	6.3e-07 (5.0e-08, 2.5e-06)	6.2e-09 (1.0e-10, 3.1e-08)
	Very low [0,5)	2.7e-03 (2.6e-03, 2.9e-03)	2.7e-03 (2.6e-03, 2.8e-03)	2.6e-03 (2.4e-03, 2.7e-03)	2.3e-03 (2.2e-03, 2.4e-03)
Education	College or university	7.9e-04 (7.2e-04, 8.7e-04)	7.8e-04 (7.1e-04, 8.5e-04)	7.7e-04 (7.0e-04, 8.4e-04)	7.1e-04 (6.5e-04, 7.7e-04)
	Elementary or less	5.2e-04 (4.0e-04, 6.7e-04)	6.5e-04 (5.0e-04, 8.1e-04)	6.6e-04 (5.0e-04, 8.3e-04)	5.0e-04 (3.8e-04, 6.3e-04)
	High school	6.3e-04 (5.8e-04, 6.8e-04)	6.4e-04 (6.0e-04, 6.9e-04)	5.5e-04 (5.0e-04, 6.0e-04)	5.0e-04 (4.7e-04, 5.4e-04)
Origin	Hispanic	6.1e-04 (5.3e-04, 7.1e-04)	7.0e-04 (6.0e-04, 8.0e-04)	7.0e-04 (6.2e-04, 7.9e-04)	7.0e-04 (6.2e-04, 7.8e-04)
	Non-Hispanic Black	1.1e-03 (1.0e-03, 1.2e-03)	9.9e-04 (9.0e-04, 1.1e-03)	8.2e-04 (7.4e-04, 9.0e-04)	7.4e-04 (6.7e-04, 8.1e-04)
	Non-Hispanic White	5.6e-04 (5.2e-04, 6.1e-04)	6.0e-04 (5.5e-04, 6.4e-04)	5.5e-04 (5.1e-04, 6.1e-04)	5.1e-04 (4.7e-04, 5.5e-04)
	Other	5.5e-04 (4.0e-04, 7.3e-04)	4.4e-04 (3.3e-04, 5.7e-04)	6.3e-04 (4.8e-04, 8.0e-04)	4.2e-04 (3.2e-04, 5.4e-04)
Prematurity	Extremely preterm ¡28w	1.7e-03 (1.6e-03, 1.8e-03)	1.5e-03 (1.4e-03, 1.6e-03)	1.5e-03 (1.4e-03, 1.6e-03)	1.3e-03 (1.2e-03, 1.3e-03)
	Moderate to late preterm [32,37)w	5.7e-04 (4.4e-04, 7.3e-04)	8.2e-04 (6.6e-04, 1.0e-03)	5.2e-04 (4.1e-04, 6.4e-04)	6.2e-04 (4.9e-04, 7.6e-04)
	Term 37w+	1.3e-04 (8.8e-05, 1.8e-04)	1.3e-04 (9.2e-05, 1.8e-04)	9.6e-05 (6.7e-05, 1.3e-04)	6.1e-05 (4.2e-05, 8.3e-05)
	Very preterm [28,32)w	7.8e-04 (5.7e-04, 1.0e-03)	1.0e-03 (7.7e-04, 1.3e-03)	5.6e-04 (4.4e-04, 7.0e-04)	7.2e-04 (5.8e-04, 8.9e-04)

References

Aitkin M, Clayton D (1980) The fitting of exponential, Weibull and extreme value distributions to complex censored survival data using GLIM. Journal of the Royal Statistical Society C 29(2):156–163, DOI 10.2307/2986301

de Beer J, Janssen F (2016) A new parametric model to assess delay and compression of mortality. Population Health Metrics 14(1):46, DOI 10.1186/s12963-016-0113-1

Berrut S, Pouillard V, Richmond P, Roehner BM (2016) Deciphering infant mortality. Physica A: Statistical Mechanics and its Applications 463:400–426, DOI 10.1016/j.physa.2016.07.031, 1603.04007

Bourgeois-Pichat J (1951) La mesure de la mortalité infantile. II. les causes de décès. Population 6(3):459–480, DOI 10.2307/1523958

Brillinger DR (1961) A justification of some common laws of mortality. In: Transactions of the Society of Actuaries, vol XIII, pp 115–119

Camarda CG, Eilers PH, Gampe J (2016) Sums of smooth exponentials to decompose complex series of counts. Statistical Modelling 16(4):279–296, DOI 10.1177/1471082X16641796

Cha JH, Finkelstein M (2016) Justifying the Gompertz curve of mortality via the generalized Polya process of shocks. Theoretical Population Biology 109:54–62, DOI 10.1016/j.tpb.2016.03.001

Choe MK (1981) Fitting the age pattern of infant and child mortality with the Weibull survival distribution. In: Asian and Pacific census forum, vol 7, pp 10–13

Clayton D, Cuzick J (1985) Multivariate generalizations of the proportional hazards model. Journal of the Royal Statistical Society A 148(2):82–108, DOI 10.2307/2981943

Clayton DG (1983) Fitting a general family of failure-time distributions using GLIM. Applied Statistics 32(2):102–109, DOI 10.2307/2347288

Cox DR (1972) Regression models and life-tables. Journal of the Royal Statistical Society B 34(2):187–220

Currie ID (2016) On fitting generalized linear and non-linear models of mortality. Scandinavian Actuarial Journal 2016(4):356–383, DOI 10.1080/03461238.2014.928230

Eilers PHC, Marx BD (1996) Flexible smoothing with B-splines and penalties. Statistical Science 11(2):89–102

Finkelstein MS (2005) Lifesaving explains mortality decline with time. Mathematical Biosciences 196(2):187–197, DOI 10.1016/j.mbs.2005.04.004

Gage TB, Dyke B (1986) Parameterizing abridged mortality tables: The Siler three-component hazard model. Human Biology pp 275–291

Heligman L, Pollard JH (1980) The age pattern of mortality. Journal of the Institute of Actuaries 107(1):49–80, DOI 10.1017/s0020268100040257

Hougaard P (1984) Life table methods for heterogeneous populations: Distributions describing the heterogeneity. Biometrika 71(1):75–83, DOI 10.1093/biomet/71.1.75

Hougaard P (1995) Frailty models for survival data. Lifetime Data Analysis 1(3):255–273

Huck P (1995) Infant mortality and living standards of English workers during the industrial revolution. The Journal of Economic History 55(3):528–550, DOI 10.1017/s0022050700041620

Knodel J, Kintner H (1977) The impact of breast feeding patterns on the biometric analysis of infant mortality. Demography 14(4):391–409, DOI 10.2307/2060586

Lehman EH (1963) Shapes, moments and estimators of the weibull distribution. IEEE Transactions on Reliability 12(3):32–38, DOI 10.1109/tr.1963.5218214

Levitis DA (2011) Before senescence: the evolutionary demography of ontogenesis. Proceedings of the Royal Society B 278(1707):801–809, DOI 10.1098/rspb.2010.2190

Lomax KS (1954) Business failures: Another example of the analysis of failure data. Journal of the American Statistical Association 49(268):847–852

Marshall AW, Olkin I (2007) Life Distributions. Springer

Mode CJ, Busby RC (1982) An eight-parameter model of human mortality—the single decrement case. Bulletin of Mathematical Biology 44(5):647–659, DOI 10.1007/bf02462273

Murphy SA, Van der Vaart AW (2000) On profile likelihood. Journal of the American Statistical Association 95(450):449–465

National Center for Health Statistics (2016) Birth cohort linked birth-infant death data files (U.S. data). URL `ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/DVS/cohortlinkedus/`

Oppermann (1870) On the graduation of life tables, with special application to the rate of mortality in infancy and childhood. Insurance Record

Prékopa A (1958) On secondary processes generated by a random point distribution of Poisson type. Annales Univ Sci Budapest de Eötvös Nom Sectio Math 1:153–170

Rogers A, Little JS (1994) Parameterizing age patterns of demographic rates with the multiexponential model schedule. Mathematical Population Studies 4(3):175–195, DOI 10.1080/08898489409525372

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

Steffensen JF (1930) Infantile mortality from an actuarial point of view. Scandinavian Actuarial Journal 1930(2):272–286, DOI 10.1080/03461238.1930.10416902

Strehler BL, Mildvan AS (1960) General theory of mortality and aging. Science 132(3418):14–21, DOI 10.1126/science.132.3418.14

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(5):313–329, DOI 10.1017/s2046167400043688

Thompson R, Baker RJ (1981) Composite link functions in Generalized Linear Models. Journal of the Royal Statistical Society C 30(2):125–131, DOI 10.2307/2346381

Vaupel JW, Yashin AI (1983) The deviant dynamics of death in heterogeneous populations. Tech. rep., International Institute for Applied Systems Analysis, URL `http://user.demogr.mpg.de/jwv/pdf/IIASA-83-001.pdf`

Vaupel JW, Yashin AI (1985) Heterogeneity's ruses: Some surprising effects of selection on population dynamics. The American Statistician 39(3):176–185, DOI 10.2307/2683925

Vaupel JW, Zhang Z (2010) Attrition in heterogeneous cohorts. Demographic Research 23(26):737–748, DOI 10.4054/DemRes.2010.23.26

Vaupel JW, Manton KG, Stallard E (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16(3):439–54, DOI 10.2307/2061224

Wienke A (2011) Frailty Models in Survival Analysis. Biostatistics Series, Chapmann and Hall, DOI 10.1111/j.1541-0420.2012.01769.x

Wittstein T, Bumsted DA (1883) The mathematical law of mortality. Journal of the Institute of Actuaries and Assurance Magazine 24(3):153–173, DOI 10.1017/s0020268100006260

World Health Organization (2006) Neonatal and Perinatal Mortality: Country, Regional and Global Estimates. World Health Organization

World Health Organization (2015) Health in 2015. From Millennium Development Goals to Sustainable Development Goals. World Health Organization