

Applying parametric models to survival data: tradeoffs between statistical significance, biological plausibility, and common sense

Alexey Golubev · Andrei Panchenko · Vladimir Anisimov

Received: 12 February 2018 / Accepted: 30 May 2018 / Published online: 4 June 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract Parametric models for survival data help to differentiate aging from other lifespan determinants. However, such inferences suffer from small sizes of experimental animal samples and variable animals handling by different labs. We analyzed control data from a single laboratory where interventions in murine lifespan were studied over decades. The minimal Gompertz model (GM) was found to perform best with most murine strains. However, when several control datasets related to a particular strain are fitted to GM, strikingly rigid interdependencies between GM parameters emerge, consistent with the Strehler-Mildvan correlation (SMC). SMC emerges even when survival patterns do not conform to GM, as with cancer-prone HER2/neu mice, which die at a log-normally distributed age. Numerical experiments show that SMC includes an artifact whose magnitude depends on dataset deviation from conformance to GM irrespectively of the noisiness of small datasets, another contributor to SMC. Still another contributor to SMC is the compensation effect of mortality (CEM): a real tradeoff between the physiological factors responsible for initial vitality and the rate of its decline. To avoid misinterpretations, we advise checking experimental results against a SMC based on historical controls or on subgroups

obtained by randomization of control animals. An apparent acceleration of aging associated with a decrease in the initial mortality is invalid if it is not greater than SMC suggests. This approach applied to published data suggests that the effects of calorie restriction and of drugs believed to mimic it are different. SMC and CEM relevance to human survival patterns is discussed.

Keywords Survival · Mortality; aging · Parametric model · Gompertz-Makeham law · Strehler-Mildvan correlation; theories of aging

...we are suffering from a plethora of surmise, conjecture and hypothesis. The difficulty is to detach the framework of fact—of absolute undeniable fact—from the embellishments of theorists and reporters.

Sherlock Holmes
Silver Blaze

Introduction

Using laboratory rodents to test interventions intended to increase healthy lifespan in humans is an essential step from basic research to its applications. It is increasingly recognized that when only the mean,

A. Golubev (✉) · A. Panchenko · V. Anisimov
N.N. Petrov Research Institute of Oncology, Pesochny-2,
Saint-Petersburg 197758, Russia
e-mail: lxglbv@rambler.ru

median or modal lifespans are indicated in a report on such a test, important information, which is present in complete survival data, is missed. Being extracted from survival curves, this information may help to define which lifespan determinants, such as aging rate, are implicated in the observed effects (Shouman and Witten 1995; Pletcher et al. 2000; de Magalhaes et al. 2005).

Under protected laboratory conditions, where predators, infections and other wildlife factors of age-independent mortality are minimized, the main determinants of lifespan include the initial viability of animals and its general decline observed with increasing age and manifested as increasing mortality rate. When a particular pathologic process is recognized as a significant contributor to the observed mortality and survival patterns, as it may be with transgenic animals deliberately made highly prone to a pathology, such as cancer, the rate of its development, which culminates in death, should be taken into account too.

Thus, two contrasting views on relationships between age-dependent mortality and lifespan may be distinguished. One of the views implies that age-dependent mortality patterns are generated by physiological changes that gradually increase chances to die of any causes, and the other view implies that physiological changes lead to death caused by a particular factor at a certain age, which varies, as all quantitative biological parameters do, according to a certain distribution.

The basic parametric models representing the two views are shown in Table 1. In demography, many other models, with much more parameters, have been suggested (de Beer and Janssen 2016; Tabeau 2001) just to ensure possibly more accurate fits to data derived from huge samples represented by human populations, without caring much about possible biological implications of the models.

The Gompertz model (GM) and its elaborations, such as Gompertz-Makeham model (GMM) [see (Olshansky and Carnes 1997; Golubev 2009; Kirkwood 2015;)] and logistic Gompertz model (LGM) (Pletcher et al. 2000), are derived from the assumption that the risk of death is a certain function of age, so that the mortality rate (hazard function) μ depends on age according to the equation $\mu(t) = \lambda \cdot e^{\gamma \cdot t}$, which is much simpler and intuitively comprehensible than the respective cumulative distribution function (CDF) and probability density function (PDF)—see Table 1.

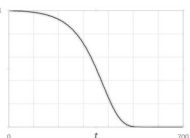
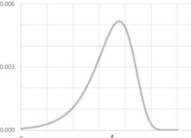

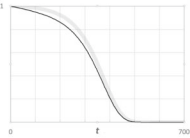
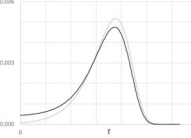
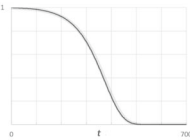
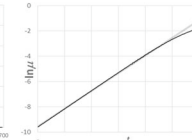
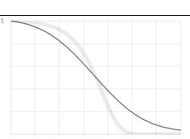
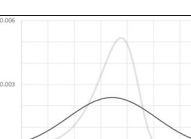
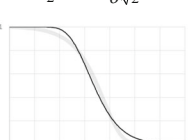
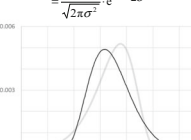
The term γ is interpreted as capturing the rate of aging, and the term λ (μ at $t = 0$), as capturing the initial vulnerability to the causes of death (the inverse of the initial robustness, vigor, or viability). By contrast, the Gaussian and lognormal models are derived from assumptions about age-at-death distributions (PDF), which in these cases are simpler and better tractable analytically than the respective hazard functions and CDFs.

GM corresponds to the view that aging causes a gradual constantly accelerating (i.e., close to an exponential) increase in the probability of death during a defined time interval. By contrast, the Gaussian and lognormal models imply that aging leads to death at a certain age \pm some random deviations; therefore, the rate of aging may be thought of as the inverse of the time to death. The Gaussian implies that the variability of ages at death is generated by numerous additive influences, whereas the lognormal implies that the effects of the influences are multiplicative. The Gaussian model is, strictly speaking, inherently inadequate to most of biological situations because the domain of its PDF extends to negative values. Although the central limit theorem justifies its applicability, as a practically acceptable approximation, to apparently symmetric distributions, whose means are positive and much greater than standard errors, the lognormal distribution has been repeatedly suggested to be used for characterization of biological variabilities, in particular, times to death of defined diseases, including cancers (Chapman et al. 2013; Limpert and Stahel 2011, 2017; Royston 2001).

The graphical inserts in Table 1 show the characteristic features of the above distributions. The plots in the inserts were built by approximating data on the survival of female NMRI mice (see below) with respective CDFs. The pure GM is distinctive in that its hazard functions is linear on a semilogarithmic scale, and its PDF is left-skewed. The lognormal PDF is right-skewed, at marked difference from the GM PDF. The plots of the hazard functions of the Gaussian and lognormal models may, if they are interpreted in GM terms, suggest that increases in the rate of aging decelerate at later ages. Noteworthy, the survival (CDF) plots, which are the ones reported usually, may be virtually indistinguishable at a glance.

The common approach to presenting survival data is to use Kaplan–Meier plots. This approach originates from pharmacological trials intended to test

Table 1 Two opposing approaches to the parametric analysis of survival and mortality data*

Variables	$N(t)$ The fraction of the members of a sample that survive to time t .	$n(t)$ The fraction of the members of a sample that die within time interval $(t, t+\Delta t)$ Age-at-death (lifespan) distribution (probability density function) PDF = $d\text{CDF}/dt$, if data are normalized to unity	$\mu(t)$ The proportion of those that die within time interval $(t, t+\Delta t)$ vs. those that survive to time t The force of mortality (hazard function = PDF/CDF)
General analytical relationships	Survival distribution (unity minus cumulative distribution function CDF , if data are normalized to unity) $= 1 - \int_0^t n(t) dt$	$= \lim_{\Delta t \rightarrow 0} \frac{N(t+\Delta t) - N(t)N(t)}{\Delta t} = \frac{dN(t)}{dt}$	$= -\frac{dN(t)}{dt} \times \frac{1}{N(t)}$
Compertz (GM)	 $= 1 - e^{-\frac{\lambda}{\gamma}(e^{\gamma \cdot t} - 1)}$	 $= \lambda \cdot e^{-\frac{\lambda}{\gamma}(e^{\gamma \cdot t} - 1)}$	 $= \lambda \cdot e^{\gamma \cdot t}$
Models based on a hazard functions**	Compertz-Makeham (GMM)	 Captures age-independent mortality.	 $= \lambda \cdot e^{\gamma \cdot t} + C$
	Logistic Compertz (LGM)	 Captures the heterogeneity of λ and/or the late-life deceleration of increases in μ	 $= \lambda \cdot e^{\gamma \cdot t} \frac{1}{1 + s \frac{\lambda}{\gamma}(e^{\gamma \cdot t} - 1)}$
Models based on a probability density functions	Normal (Gaussian) (NM)	 $= 1 - \frac{1}{2} \cdot [1 + \text{erf}(\frac{t-S}{\sigma\sqrt{2}})]$ ***	 $= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(t-S)^2}{2\sigma^2}}$ ****
	Lognormal (LNM)	 $= \frac{1}{2} \cdot [1 + \text{erf}(\frac{\ln(t)-S}{\sigma\sqrt{2}})]$ ***	 $= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{[(\ln(t)-S)]^2}{2\sigma^2}}$ ****

* Thick faint lines on graphical inserts, which show plots derived from normalized data on NMRI mice survival (see below), relate each time to GM

** The parameters λ , γ , and C are usually interpreted, respectively, as the initial mortality (the inverse of the initial vitality), the rate of aging, and the age-independent (background, external) mortality, and s , as quantifying the heterogeneity of a population (cohort)

*** erf is the error function, which is not representable with elementary functions; only numerical treatment is possible

**** The parameterized expressions for respective PDF to CDF ratios are omitted because of their awkwardness

interventions in a specific pathology and originally is designed to exclude death cases unrelated to a pathology of interest [see (Breslow 1992)]. The exclusion procedures, which are known as censoring, may be irrelevant in experimental gerontology when a cohort under study is followed up to the death of its last member, and death cases are not differentiated by their causes. Therefore, presenting survival data as points that show the numbers (fractions) of survivors versus time may be admitted as adequate.

However, in studies with mice, the sizes of experimental groups are usually not enough to provide in each particular case for a reliable judgment, even in qualitative terms, about the rule that governs the time course of survival in the respective parent population and generates the corresponding age-at-death distribution, which may be described with a parametric model, in particular, interpretable in terms of the initial viability and the rate of its decline. This is even more so with regard to the quantitative estimates of the parameters of any model chosen.

The factors that have been shown by different authors (Witten and Satzer 1992; Eakin et al. 1995; Shouman and Witten 1995; Pletcher 1999; Pletcher et al. 2000; Golubev 2004, 2009; Yen et al. 2008; Petrascheck and Miller 2017; Bokov et al. 2017; Tarkhov et al. 2017; Tai and Noymer 2018) to bias the estimates of the parameters of any analytical model, such as a GM-based one, which is used to approximate survival data, include the following:

- Differences in the sizes of study samples drawn from a parent population, which make parameter estimates for each sample different. In particular, increasing the initial size of a study cohort will tend to decrease the estimates of λ and increase the estimates of γ in GM.
- The accuracy of age-at-death records, which depends on time intervals chosen to check for the presence of death animals and thus on the discretization of the survival time of a cohort.
- The positive correlation between the survival time of a cohort and its initial size.
- The mathematical techniques or algorithms chosen to approximate a limited set of discrete data with a smooth function.
- Uncontrolled environmental changes that occur through the time of observation of a cohort.
- The inhomogeneity of a cohort under study.

- Deviations of the physiological decline from linearity, especially at the extremes of cohort lifespan.
- The assumption that the values of some of the essential parameters of a model, e.g. the background mortality C in GMM, are invariably too small to matter in practical fitting terms.
- Using of period data to make inferences about cohorts.

Importantly, with a two- or three-parametric model, biases simultaneously introduced to the estimates of its parameters by one of the above factors may introduce an apparent correlation between the parameters that otherwise must be independent according to the basic assumptions of the model, which thus may become compromised.

One way to obviate at least some of the above issues is to increase the initial sizes of experimental cohorts, e.g., it has been shown that cohort size-dependent biases of the estimates of GM parameters become negligible at initial cohort sizes above 160 (Bokov et al. 2017). However, this approach will not work when the factors that are included in the second half of the above list are at work. Besides, taking this approach straightforwardly will invalidate most of experimental findings obtained so far with vertebrates. A similar problem with clinical trials resulted in the development of meta-analysis procedures, which were recently adopted to analyses of published results of gerontologic experiments (Nakagawa et al. 2012; Simons et al. 2013; Liang et al. 2018). Another approach is to use the data presented in published works as tables or plots (survival curves) for treating them all in the same way, which may be not the one exercised in all of the source publications or in any of them. Several analyses like that were performed using different approaches to finding and comparing the parameters of models chosen to fit experimental data (de Magalhaes et al. 2005; Hughes and Hekimi 2016; Shen et al. 2017; Simons et al. 2013; Yen et al. 2008). Each time a single approach is used, its possible biases are systematic across all datasets and, therefore, the resulting unidirectional inaccuracies hardly interfere in a significant manner with inferences from comparing changes in a parameter of choice across different experiments.

DeMagalhaes et al. (2005) reported on their analysis of the results of two dozen studies where genetic

modifications influenced lifespan in mice. GM parameters were derived from the estimates of hazard, which were obtained by relating death events to appropriately chosen age intervals. The log-transformed numerical values were then fitted with a straight line using linear regression. It was concluded that for the comparison of experiments involving small numbers of animals the basic GM, which does not account for the background mortality (which is captured by the Makeham parameter C) and of the heterogeneity of animals, is suitable. It was shown that, in many cases, changes in lifespan claimed by the authors of the original findings to suggest changes in the rate of aging (γ according to the notation used in the present paper) were more consistent with changes in the initial mortality (λ). Remarkably, for SAMP (senescence acceleration prone) mice, which are generally believed to exhibit an accelerated aging, plotting of log hazard against age, which according to GM must yield a straight line, yielded patterns qualified as “bizarre” (de Magalhaes et al. 2005).

Another remarkable observation mentioned by deMagalhaes et al. is a negative linear correlation between γ and $\ln\lambda$. This correlation was predicted by B. Strehler and A. Mildvan based on their general theory of aging and confirmed in the same paper (Strehler and Mildvan 1960) by comparing Gompertz parameters, which were derived from period data, across human populations in different countries. These observations were later reproduced many times by other authors [e.g. (Anderson et al. 2017; Finkelstein 2012; Gavrilov and Gavrilova 1991; Zheng et al. 2012)], but their significance was disputed. At present, opinions on the Strehler-Mildvan correlation (SMC) range from attributing it to artifacts of the mathematical treatment of survival and/or mortality data (Burger and Missov 2016; Tarkhov et al. 2017) to providing reasons to believe that the artifact may be superimposed on manifestations of the real heterogeneity of a study population (Avraam et al. 2016; Shen et al. 2017; Zheng et al. 2012) and/or the basic physiological tradeoffs relevant to aging (Golubev 2009; Golubev et al. 2017b; Shen et al. 2017).

Yen et al. (2008) compared the approach used by DeMagalhaes et al. (2004) with extracting GM parameters from survival curves by fitting them to GM CDF using nonlinear regression or maximum likelihood estimation (MLE) to find the combinations of parameters that could fit experimental datasets best.

Yen et al. confirmed the conclusion (Eakin et al. 1995; Pletcher 1999) that MLE is superior to other approaches to survival curve fitting. The authors also confirmed that increases in lifespan are not necessarily associated with decreases in the rate of aging. They did not check GM parameters for a correlation; however, SMC can be revealed by plotting their estimates of γ against $\ln\lambda$.

Hughes and Hekimi (2016) used MLE to extract λ and γ values from survival data and found correlations between median lifespans with λ but not γ in a series of murine strains. The authors did not mention SMC. The correlation however is apparent in supplementary materials to their paper.

Thus, extracting GM parameters from survival curves may provide for meaningful conclusions when they are based on comparisons across a set of similarly treated different samples even if each sample is too small for drawing reliable inferences from it. The use of an elaboration, such as GMM or LGM, of the basic GM in such cases adds little if anything because the subtle features that are captured by additional parameters may be revealed only when samples are big enough to provide for a sufficient resolution (Bokov et al. 2017). Anyway, the prerequisite of such analysis is the in-principle conformance of all samples to GM, either basic or appropriately elaborated. This may not be so with genetically modified mice, which often develop a dominant pathology resulting in death rather than exhibit a gradual increase in vulnerability to many causes of death. Another open question is the significance of correlations between GM parameters.

To address these issues we used a unique combination of survival datasets obtained in a single laboratory over years of testing different interventions in the lifespans of mice of several strains kept under similar standard conditions (Anisimov et al. 2013). Most of experiments at the base of the present study have been published earlier (references will be given at appropriate places in Results and Discussion). Because samples from a single stock were used as controls in several experiments performed over a decade, the use of the control datasets allows comparisons of samples attributable to a single parent population, as well as comparisons across strains and experimental conditions.

To determine in an unbiased manner (that is without taking into account the possible interpretations of the parameters of models under comparison),

which CDF is most relevant to a given dataset and across datasets, we tried the approach found to be helpful in comparing different PDF models applied to cell cycle time variability (Golubev 2016) using a curve-fitting tool specifically designed for such comparisons: TableCurve 2D (TC2D) ver. 5.1 (Systat Software Inc., San Jose, CA, fully functional trial version available at <http://www.sigmaplot.com/products/tablecurve2d>).

Methods and materials

Numerical data were derived from published survival curves using the CurveSnap free digitizer application (<https://curvesnap.en.softonic.com/>) upon assigning unity to the tops of the original plots. With Kaplan–Meier type plots, only the lower points of their vertical segments (the proportions of mice survived up to defined time points) were recorded.

The numerical data extracted from survival plots were treated with TableCurve 2D (TC2D) ver. 5.1 (Systat Software Inc., San Jose, CA) where nonlinear maximum likelihood estimation (MLE) is employed to compare its inbuilt as well as user-defined functions (UDF). The results are presented as ranked lists of functions able to fit an input dataset, each function associated with the estimates of its fit (determination coefficient r^2 , DOF-adjusted r^2 , and f-statistic values) and of its parameters and respective standard errors and confidence intervals. The plots of the analytical approximations of datasets are displayed for each function, and their numerical representations may be used for further analysis. Among the inbuilt functions, 15 (transition functions) are applicable, in principle, to survival data. They include, in particular, CDFs corresponding to the normal, lognormal, and Weibull distributions, and several versions of logistic dose–response and sigmoid functions. The functions that feature symmetric PDFs, which extend to the negative domain, were disregarded as having no physical sense in the present context, except for the cumulative normal distribution, which was used for comparative purposes.

A controversial issue is the choice of the initial age. In many experiments and publications, no deaths were observed or reported until the age of 100–200 days, and no points are present in these intervals. We reasoned that upon the assumption that a sample under

consideration conforms to a parametrically representable law, which governs the survival pattern of the respective parent population, it is possible to calculate the proportion of death cases that must fall within the period where no actual deaths were observed. The calculated proportions are such that sample sizes required for the counts of deaths occurring before 100 days to exceed unity must be too big for routine practice. For example, with 129/Sv mice, GM suggests that one has to start with thousands of mice or to perform hundreds of experiments with samples of conventional sizes about 50 (this does not rule out the possibility of an accidental observation of death in even a smaller-scale single experiment). A similar problem, by the way, relates to the opposite end of an analytical approximation of a survival curve. Even if the approximation extends to infinity, this does not mean that (murine) lifespan is infinite: with any limited cohort, its actual survival discontinues near a time point where the number of survivors reaches unity. This means that the time of the death-caused exhaustion of a cohort must increase with increasing its initial size. The fact that the maximal observed lifespans correlate with the sizes of the populations of a species studied is long acknowledged (Wilmoth et al. 2000).

Another issue is whether the day of birth or the day of reaching the mature state should be taken as the starting point for analysis. Clearly, survival patterns before and after reaching the maturity must be different. However, with samples as small as those being dealt with, both the maturation and the post-maturation periods, until at least 100 days of age, are similar in showing zero mortality. With account of the lack of a universal agreement concerning the age of maturation in mice (the onset of estrous cycles, or body weight constancy, or established neuronal connectivity?) it was decided to start analysis from day “zero”. Systematic biases in parameter estimates when they all are based on a single assumption must be irrelevant for comparisons between samples. For example, for all samples assumed to conform to GM, shifting the starting point from 0 to 100 or 200 days will produce virtually no effect on the estimates of γ and will decrease the estimates of λ proportionally about 3 or 9 times, respectively, which will make no impact on comparing λ across samples.

As an example, treating the data (Anisimov et al. 2011b) on the survival of 155 female SHR mice (Swiss

H-derived outbred mice originating from Rappolovo Farms) with TC2D tool will be described below.

The list of functions ranked by their fits (determination coefficients r^2) to survival data is presented in Table 2. All user-defined functions (UDF) are printed in bold type, and the only user-defined function (i.e., Gompertz CDF) that does not have inbuilt analogues is printed in italics. All inbuilt functions have an additional adjustable scale parameter designated as H here because it captures transition height. Since all input data were normalized to unity, the inbuilt functions were supplemented with UDFs where H was deleted. The physical (biological) meanings of the parameters γ , λ , S , and σ are assumed to be intuitively clear (see Table 1). No effort to interpret the parameters designated as a and b was made. The fits of the functions having ranks from eighth to thirteenth are overly inferior to those of the functions having higher ranks. The same is true for UDFs of cumulative gamma and Wald (inverse Gaussian) distributions. These observations are relevant to all datasets that were analyzed.

The models, all defined with respective UDFs and constrained to unity at time 0, that were eventually chosen to compare them through the present paper are presented in bold font. Their rankings may be different in other cases. However, with r^2 values as high and differences between them as small as shown in Table 2, any model based on some premises applied to a single dataset may be construed as suggesting that the premises are valid, if no comparison with competing models is made. Moreover, a single

comparison is insufficient when differences between fits are as small as in the present case. An unbiased choice of the most appropriate model should be based on its ability to outperform other models systematically.

Results and discussion

Comparing the performance of parametric models applied to murine survival datasets

Table 3 compares the performances of four models highlighted in Table 2 upon their application to survival curves of non-genetically modified female mice of several strains. The superiority of GM, which is the only model whose PDF is skewed negatively, is small if judged by r^2 ; however it is quite consistent ($p < 0.01$ even by the simplest nonparametric criterion, i.e. sign test (Dixon and Mood 1946)).

The superiority of GM is not simply because it is more flexible in comparison with competitors. The results of its application are far less consistent when data on transgenic HER2/neu female mice survival are treated in the same way as above. The mice have a copy of rat HER2 gene coding for an epidermal growth factor receptor. Because of that, they almost invariably die of mammary cancer, the mean lifespan being three times shorter than that of their strain of origin, FVB/N.

The basic differences in survival patterns between HER2/neu and FVB/N mice compared in the same

Table 2 A ranked list of transition function applied to data (Anisimov et al. 2011a, b) on the survival of female outbred Swiss-derived SHR mice*

**Italic*—user-defined function having no inbuilt analogues
***Bold** characters—functions chosen for being used in analysis

Rank	Determination coefficient r^2	Function
1	0.995	<i>Gompertz cumulative</i> (γ, λ)
2	0.993	Gaussian cumulative (H, S, σ)
3	0.988	Gaussian cumulative (S, σ)
4	0.987	Logistic dose–response_ (H, a, b)
5	0.986	Lognormal cumulative (H, S, σ)
6	0.968	Logistic dose-response (a, b)
7	0.966	Lognormal cumulative (a, b)
8	1e–08	Asymmetric sigmoid_ (H, a, b)
9	1e–08	Cascade_ (H, a, b, c)
10	1e–08	Exponentially modified Gaussian_ (H, a, b, c)
11	1e–08	Pulse-power cumulative_ (H, a, b, c)
12	1e–08	Weibull cumulative (H, a, b, c)
13	1e–08	Asymmetric sigmoid reverse (H, a, b, c)

Table 3 The comparative performance of parametric models applied to control data on non-genetically modified female mice survival

Strain	n	Mean age	Data source	Model (the skewness of its PDF)*							
				GM (-)		NM (0)		LDR (+)		LNM (++)	
				r ²	Rank	r ²	Rank	r ²	Rank	r ²	Rank
129/Sv	48	711 ± 24	Anisimov et al. (2010b)	.982	1	.975	2	.963	4	.964	3
129/Sv	65	786 ± 19	Anisimov et al. (2011b)	.991	1	.972	2	.965	3	.964	4
129/Sv	46	763 ± 22	Popovich et al. (2013)	.992	1	.982	2	.972	3	.971	4
129/Sv	35	758 ± 32	Anisimov et al. (2015)	.978	1	.969	2	.938	3	.936	4
SHR [#]	80	559 ± 23	Anisimov et al. (2011b)	.995	1	.989	2	.968	3	.966	4
SHR [#]	119	511 ± 20	Anisimov et al. (2011a)	.996	1	.995	2	.977	3	.976	4
SHR [#]	60	596 ± 16	Anisimov et al. (2016)	.993	1	.987	2	.977	3	.976	4
NMRI	50	346 ± 12	Popovich et al. (2004)	.994	1	.989	2	.979	3	.978	4
FVB/N	50	600 ± 12	Panchenko et al. (2016)	.984	1	.969	2	.938	3	.934	4

[#]Outbred mice derived from the Swiss H strain

*(-) PDF is skewed negatively; (0) PDF is symmetrical; (+) PDF is skewed positively; (++) PDF is strongly skewed positively
The bold characters indicate only that the respective parameters are the most significant

experiment (Panchenko et al. 2016) are shown in Figs. 1 and 3.

At a glance, the approximations of HER2/neu survival patterns with the three functions chosen for being compared are virtually identical. A preference for any of them will be precarious if based on

negligible differences between r² values. Notably, the preference for GM to be applied to both of the survival plots would suggest an accelerated aging in HER2/neu mice.

An attempt to define preferences based on the systematic comparative over- or underperformance of

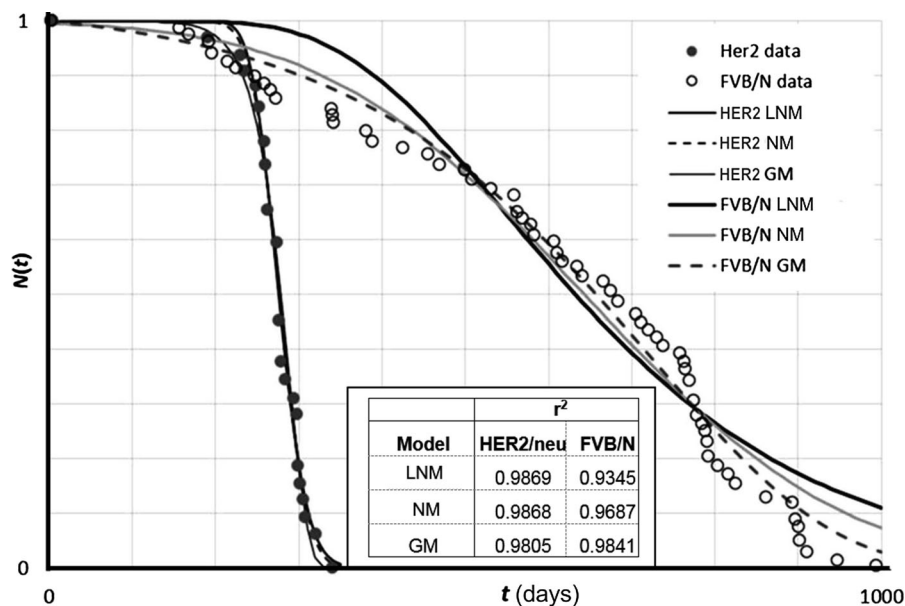


Fig. 1 Survival patterns of female HER2/neu and FVB/N mice

a model (Table 4) leads to inconsistent results for HER2/neu mice, although suggests that the functions that feature non-skewed (normal model, NM) and/or slightly positively skewed PDFs (logistic dose response model, LDRM) are somewhat advantageous.

Statistical reasons to prefer this or that model may be supplemented with the biological plausibility of the estimates of its parameters reinforced with common sense. Figure 2 shows the estimates of λ derived from GM applied to murine survival data and to data (Golubev 2009) derived from human mortality analysis.

The estimates of λ suggest, contrary to common sense, that the initial viability of transgenic cancer-prone HER2/neu mice is higher than that of their strain of origin as well as of other genetically non-modified strains. Moreover, HER2/neu λ values overlap with λ values typical for humans, which might suggest that HER2/neu mice, unlike other murine strains, may be initially as death-proof as humans are. With this in mind, it is reasonable in this case to reject GM in favor of either LDRM or lognormal model (LNM), which are irrelevant to the notion of aging. The latter choice is supported by LNM applicability to cancer patient survivorship as well as to latent periods of other diseases (Limpert and Stahel 2017; Royston 2001;

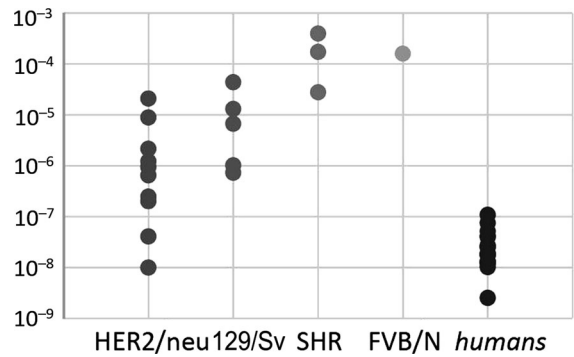


Fig. 2 The estimates of the initial mortality (λ) derived by applying GM to data on the survival of four murine strains. Human data are the same as used in (Golubev 2009)

Spratt 1969; Wang et al. 2010). It comes out in the final account that the cohorts of FVB/N, as well as 129/Sv and SHR mice, die out because of exponentially increasing chances to die, which is consistent with aging, whereas HER2/neu mice are doomed to die at a certain age \pm some relatively small random deviations unrelated to aging. The difference between the age-at-death distributions of FVB/N and HER2/neu mice are shown in Fig. 3. Note that skewness is negative in the former and positive in the latter case.

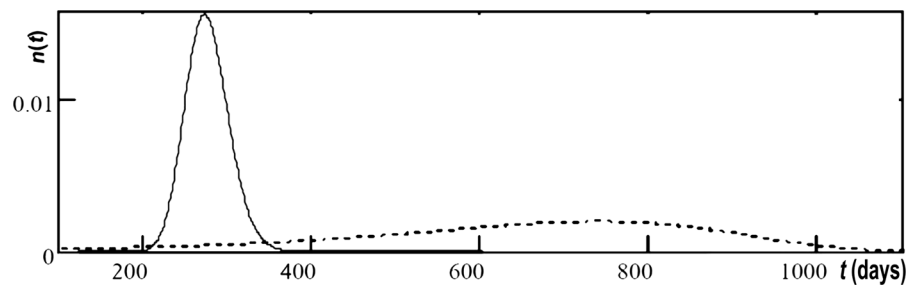
Table 4 Ranking of four models applied to control survival datasets obtained in experiments with female HER2/neu mice

Data source	Mean age	n	The ranks the fits of a model (PDF skewness)* to data			
			GM (-)	NM (0)	LDRM (+)	LNM (++)
Anisimov et al. (2003)	281 ± 8	30	1	2	3	4
Popovich et al. (2004)	321 ± 10	29	4	3	1	2
Anisimov et al. (2005)	264 ± 4	34	4	3	1	2
Anisimov et al. (2010a)	285 ± 12	15	4	1	2	3
Iurova et al. (2010)	289 ± 6	76	1	2	3	4
Anisimov et al. (2010c)	296 ± 6	28	4	3	1	2
Tyndyk et al. (2012)	292 ± 4	33	1	2	3	4
Popovich et al. (2014) [#]	268 ± 15	12	1	2	3	4
Popovich et al. (2014) [#]	278 ± 5	32	4	1	2	3
Popovich et al. (2014) [#]	285 ± 6	19	1	2	3	4
Panchenko et al. (2016)	278 ± 5	32	4	3	1	2
Unpublished (2017)	274 ± 5	55	4	1	2	3
Ranks sum			33	25	25	38

[#]Three controls used for different ages of the onset of experimental treatment

*(-) PDF is skewed negatively; (0) PDF is symmetrical; (+) PDF is skewed positively; (++) PDF is strongly skewed positively

Fig. 3 Age at death (lifespan) distributions of HER2/neu mice (solid line) and their parent FVB/N mice (dotted line) according to their best-justified lognormal and Gompertz models, respectively



One may say that it is hardly a revelation to show that the survival and mortality patterns of highly cancer-prone transgenic mice are irrelevant to aging. However, when the approach used to show this is applied to SAMP mice, which are commonly regarded as models of accelerated aging, less anticipated concerns are raised. Table 5, where the fits of different models to published data on SAMP mice survival are compared, provide reasons to doubt whether senescence is what may be really responsible for the decreased mean lifespan of SAMP mice. Even though available evidence is limited, Table 5, compared with Table 3, is far less in favor of GM and, in this regard, is more similar to Table 4.

Further evidence that may clarify the applicability of GM-based models to SAMP, HER2/neu, and non-genetically modified mice comes from correlations between γ and λ that emerge when GM is applied to all murine strains invariably.

Correlations between the parameters of models applied to murine survival datasets

Relationships between the values of γ and λ derived from GM and its elaborations applied to data on female 129/Sv mice survival are shown in Fig. 4.

A striking feature of the relationships between $\ln\lambda$ and γ is a negative linear correlation between their apparent values related to four control samples of the same murine strain. It is recognized that apparent changes in γ and λ are correlated in this way, which is generally known as the Strehler-Mildvan correlation (SMC), when pure GM is applied to datasets that are consistent with GMM and differ only in the background mortality captured by C (Gavrilov and Gavrilova 1991; Golubev 2004). However, GMM or LGM applied to control 129/Sv data does not eliminate SMC. Moreover, all correlations in Fig. 4 look as

dead rigid interdependencies, which are too good to be attributed to a biological factor and thus are likely to be artificial.

A reasonable explanation of this phenomenon is the tendency of the parameters of any non-linear multi-parametric model to become cross-correlated when it is applied to small datasets (Johnson 2000). Indeed, the parameters of the lognormal or LDR models are also cross-correlated if the models are used to fit real 129/Sv survival datasets as well as datasets generated numerically with GM or GMM, but not with the lognormal or LDR model, respectively (not shown).

SMC is conspicuous in its peculiar form: γ shows negative linear correlation with $\ln\lambda$ ($\ln\lambda = -A\gamma - B$). The origin of this peculiarity, which is rooted in the pre- and post-exponent positions of the two parameters, has been investigated based on huge experimental *C. elegans* datasets in (Tarkhov et al. 2017) and illustrated with demographic human survival data (Burger and Missov 2016; Tarkhov et al. 2017). Its manifestations in and implications for significantly smaller experimental murine datasets are explored in the present work.

To check whether SMC arises when samples under study, which are not strictly consistent with GM, feature differences in γ or λ rather than in C , numerical experiments were performed. A survival plot was generated using a GMM CDF at $\gamma = 0.0009$, $\lambda = 0.5 \times 10^{-6}$, and $C = 0.0003$, which are within the ranges found when GMM is applied to datasets explored in the present work. Then each of the parameters was altered in several steps while the other two were constant. The resulting three series of modeled plots are shown in Fig. 5.

Data points were taken from each plot in 40-day intervals to make 30 entries in each dataset, as in a typical experiment with mice. The datasets were fitted

Table 5 Comparative performance of four parametric models applied to published data on the survival of accelerated senescence-prone (SAMP) mice

Data source	Mean lifespan	n	Model (the skewness of its PDF)*								
			GM (-)		NM (0)		LDR (+)		LNM (++)		
			r ²	Rank	r ²	Rank	r ²	Rank	r ²	Rank	
Carter et al. (2005) [#]	SAMP8	≈ 400	237	.9894	4	.9981	3	.9983	2	.9985	1
	SAMP10	≈ 420	169	.9886	4	.9961	2	.9965	1	.9959	3
Takeda et al. (1981)		≈ 290	>500	.9972	2	.9992	1	.9960	3	.9964	4
Porquet et al. (2013)		≈ 330	50	.9813	4	.9890	3	.9906	2	.9913	1
(Umezawa et al. 2000) ^{##}		357 ± 21	21	.9838	1	.9793	2	.9659	4	.9661	3
		426 ± 24	20	.9861	4	.9927	1	.9871	3	.9884	2
Ranks sum					19		11		15		16

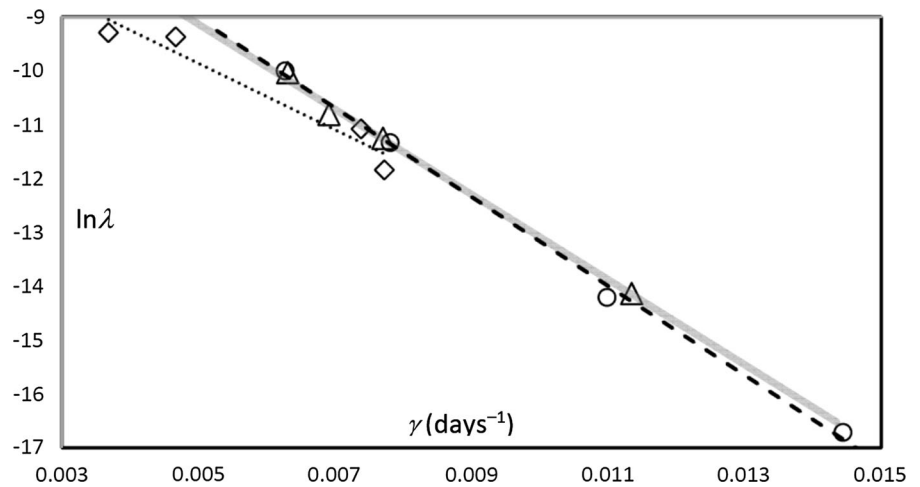
[#]Color tracking instead of point selection was used to digitize plots

^{##}Two different calorie-matched nutritional regimens

*(-) PDF is skewed negatively; (0) PDF is symmetrical; (+) PDF is skewed positively; (++) PDF is strongly skewed positively

The bold characters indicate only that the respective parameters are the most significant

Fig. 4 Correlations between the estimates of $\ln\lambda$ and γ derived from four sets of control female 129/Sv mice survival data (Table 3) according to GM (triangles, thick gray line), GMM (circles, dashed line), and LGM (diamonds, dotted line)



with GM, GMM and LGM using the TC2D tool. The resulting estimates of γ and $\ln\lambda$ are shown in Fig. 6.

An anticipated observation, which is still worthy of mentioning, is that, when a right model is used, the results of fitting shown in Fig. 6 exactly conform to the input datasets used to construct the CDFs to be fitted (Fig. 5), even though the datasets are quite small. E.g., γ varies at a constant λ (filled squares in Fig. 6) when GMM is used to fit CDF generated using GMM at variable γ and constant λ and C . However when the same CDF is fitted with GM, both γ and λ vary in a negatively correlated way (open squares in Fig. 6),

consistent with SMC, although not its original version, which implies variations in C .

A less anticipated conclusion from Fig. 6 is that, if λ is varied in GMM at constant C and γ , then GM γ and λ will be correlated positively, contrary to SMC. Numerical experiments with other models (not shown) confirm that positive, not only negative, correlations between γ and λ are possible if pure GM is applied to plots generated by varying γ or λ in a GM-based model when its additional parameter is kept constant. However, the correlations are consistently negative when

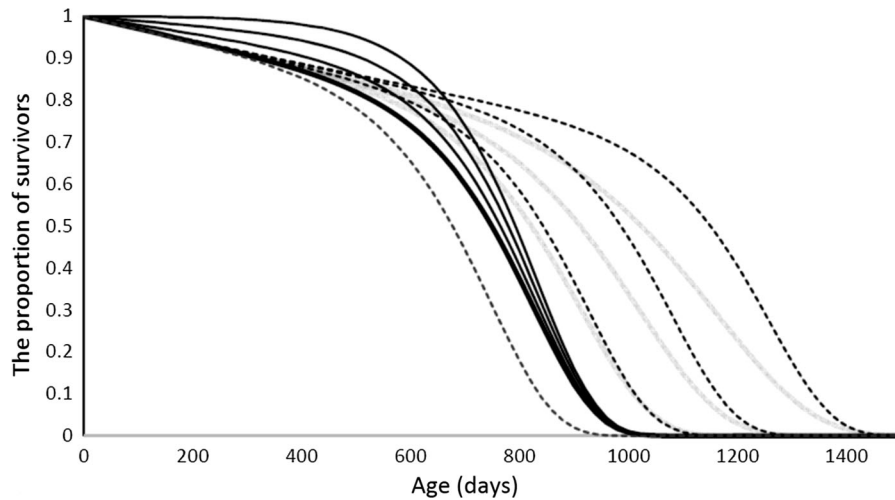


Fig. 5 Modeled GMM survival plots. Thick solid line: the initial plot ($\gamma = 0.009$, $\lambda = 0.5 \times 10^{-6}$, and $C = 0.0003$). Thin solid lines are constructed by decreasing C to 0.0002, 0.0001, and 0.0000. Thick gray lines are constructed by decreasing γ to 0.008, 0.007, and 0.006. Dotted lines are at $\lambda = 10^{-5}$ (on the left of the initial plot), and $\lambda = 2 \times 10^{-6}$, $= 10^{-6}$, and $= 5 \times 10^{-7}$ (on the right)

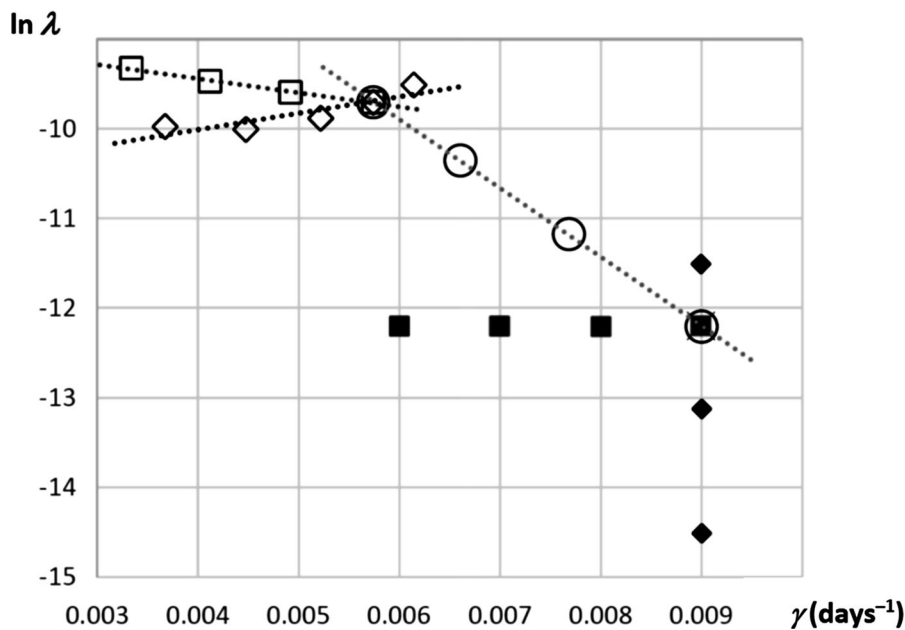


Fig. 6 Apparent changes in GM parameters (open symbols) when GM is used to fit survival plots (CDFs) constructed by varying GMM parameters (filled symbols). When only γ is varied in GMM (filled squares) at constant λ and C , then both, γ and λ , vary in GM in a *negatively* correlated manner (open squares). When only λ is varied in GMM (filled rhombs) at

constant C and γ , then both, λ and γ , vary GM in a *positively* correlated manner (open rhombs). Variation of C at constant γ and λ in GMM cannot be shown in this plot. The resulting PDFs and CDFs of GMM are shown in Fig. 5. When they are fit with GM, λ and γ both vary in a negatively correlated manner (open circles)

the additional parameter is varied or when a model unrelated to GM is used to generate input data.

With any GM-based model applied to survival/mortality data that are not strictly consistent

with it, the relationships between γ and $\ln\lambda$ are usually shaped as a linear negative correlation, which is commonly termed the Strehler-Mildvan correlation (SMC). This term will be used below until reasons will

be provided to distinguish SMC from other types of relationships between γ and $\ln\lambda$.

In particular, negative correlations between γ and λ will arise when fluctuations of experimental or demographic survival and mortality patterns are unrelated to changes in the initial viability of organisms (captured by the inverse of λ) and/or in the rate of its decline (captured by γ). Because this effect is produced by the statistical properties of fitting any noisy data with a nonlinear multiparametric model (Johnson 2000), it will be manifested irrespectively of the essential relevance of a model to the phenomenon reflected by the data.

However, even if data are not noisy at all, as it is in Fig. 6, a cross-correlation between the parameters of a nonlinear multiparametric model will still emerge if the model is not quite relevant to data. This must be true when GM is applied to data that are more consistent with GMM or GLM. This must be even more true when any GM-based model is applied to survival/mortality patterns that are generated by factors unrelated to aging, e.g. by a disease, and thus are consistent with a positively skewed age-at death distribution, such as the lognormal, which is the most reasonable model for HER2/neu mice survival data.

Figure 7 shows that an apparent rigid SMC-like relationship emerges when HER2/neu mice datasets, which are essentially inconsistent with any GM-based model (see above), are treated according to GM. Notably, the resulting pattern is quite distinct from patterns formed by the datasets that are relevant to GM, such as the pattern featured by control 129/Sv

data. The dataset related to FVB/N, the strain from which HER2/neu mice are derived, clusters with 129/Sv rather than HER2/neu data.

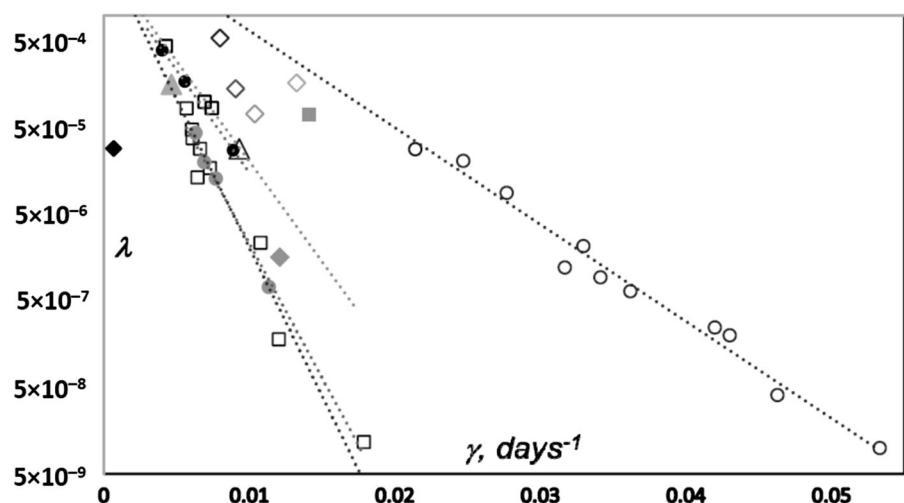
The most likely contributor to correlations between $\ln\lambda$ and γ in 129/Sv mice must be, in this case, the random deviations of the survival patterns featured by small datasets from the basic pattern featured by the parent population and captured by the respective parametric model. Indeed, this correlation is also evident when all control 129/Sv data points are combined, mixed, and then randomized into three subsamples, each fitted with GM. The resulting estimates of γ , which range from 0.0106 to 0.0114, exactly fit the $\ln\lambda$ versus γ regression shown in Figs. 4 and 7.

Similar results have been obtained with outbred female SHR mice; however, the resulting SMC is clearly different from that of inbred 129/Sv mice (Fig. 7).

The only our dataset related to another Swiss-derived outbred strain mice, NMRI, which are notably short living (Table 3), falls apart from both 129/Sv and SHR trends.

Also shown in Fig. 7 are the results derived from data on the survival of several murine strains bred at Jackson Laboratories (available at <http://www.jax.org/research-and-faculty/research-labs/the-harrison-lab/gerontology/available-data>). Remarkably, two types of datasets may be distinguished in the results: one showing the same SMC trend as found in inbred 129/Sv mice, and the other, the same as found in outbred SHR mice.

Fig. 7 SM correlations in different strains of mice. Open circles: HER2/neu; gray circles: 129/Sv; gray triangle: FVB/N; black circles: SHR; open squares: strains bred at Jackson Laboratories; gray square: NMRI; open diamonds: “senescence accelerated-prone” SAMP8; gray diamond: “senescence accelerated-resistant” SAMR; black diamond: the only SAMP8 best fitted with GM; open triangle: SAMP10



The datasets related to “senescence acceleration prone” SAMP mice (see Table 4) cluster mostly with HER2/neu rather than 129/Sv or SHR data. At the same time, data on the SAMR (“senescence acceleration resistant”) mice cluster with 129/Sv data. Among the two deviate SAMP points, one (black diamond) relates to the only SAMP8 dataset that shows the best fit to GM, and the other (gray diamond) relates to SAMP10 mice.

Taken together, the above casts doubt on the adequacy of SAMP mice as a model of accelerated senescence. It seems more likely that some dominant pathology, such as amyloidosis (Brayton et al. 2012), leads these mice to death at an age distributed according to a lognormal pattern, which may be approximated with a normal distribution featuring much greater variance than that seen in the case of HER2/neu mice. Because the lifespan of SAMP mice is greater than that of HER2/neu mice, the pathological process that terminates the lives of the former can spare more time for aging proper to be manifested. This must result in mixed survival patterns where the contributions of their main determinants vary because of differences in handling procedures used in different laboratories.

Another argument strengthening doubts concerning SAMP mice is the “bizarre” pattern of their mortality, which has been interpreted as showing a marked late-life deceleration of their initially rapid age-associated increase in mortality rate (de Magalhaes et al. 2005). The hazard function of the lognormal model clearly shows this type of behavior (see Table 1, line 4). Being interpreted as reflecting the rate of aging, this would suggest its deceleration in cases where there is no sense in speaking of aging at all, as it is the case with HER2/neu mice and is likely, according to the present analysis, to be the case with SAMP mice.

Correlations between SM parameters upon experimental modifications of murine lifespan

Other concerns raised by the above analysis relate to the interpretation of experimental changes in survival patterns even if they are reasonably attributable to a GM-based model. Given that an agenda is to ascertain whether a decrease in aging rate is responsible for the observed increase in the mean lifespan, it is important

to be sure that a decrease in γ is not attributed to an artifactual component of SMC.

The presumably non-artifactual component, which is termed the compensation effect of mortality (CEM), was initially thought of as a correlation between γ and λ that may remain when changes in the age-independent mortality, which is captured by C in GMM, are accounted for in an analysis of survival and mortality patterns (Gavrilov and Gavrilova 1991). Among the possible contributors to CEM, there may be a real physiological tradeoff between allocation of body resources (i) to the ability to withstand the causes of death and (ii) to mitigating the deterioration of this ability, i.e., to reducing the rate of aging. These physiological relationships must be manifested in a population of organisms, in which such relationships are at work, as changes in the demographic parameters λ ($1/\lambda$ reflects the initial viability) and γ (reflects the rate of aging). Hereinafter, CEM will be viewed as the physiological contributor to SMC, and SMC as a combination of CEM with all sorts of artifacts.

To get an insight into the consequences of CEM for relationships between changes in the demographic aging rate and initial mortality and the mean lifespan, several series of GM PDFs (normalized lifespan distributions) were constructed upon the condition that the relationships between the physiological determinants of the initial viability and the rate of aging are such that $\ln \lambda = -B - A \times \gamma$, where the initial B and A were chosen from ranges that correspond to SMC shown in Figs. 4 and 7. The resulting CDFs and PDFs are shown in Fig. 8.

The mean lifespans were calculated for each series of PDFs produced by varying γ within the range corresponding to the real estimates of γ for mice. This was done at the initial values of A and B and at somewhat decreased and increased values of either A or B . The calculated mean lifespan values were plotted against the input values of γ (Fig. 9a) or respective calculated values of λ (Fig. 9b).

Figure 9 suggests that if in a population (cohort) its survival pattern conforms to a GM-based model and there is a real physiological tradeoff between the viability of organisms and its age-associated decline, which is expressed at the level of populations as a negative linear correlation between γ and $\ln \lambda$, then any increase in the mean lifespan must be associated with a primary decrease in the initial mortality (λ) and the associated increase in the rate of aging (γ). Although

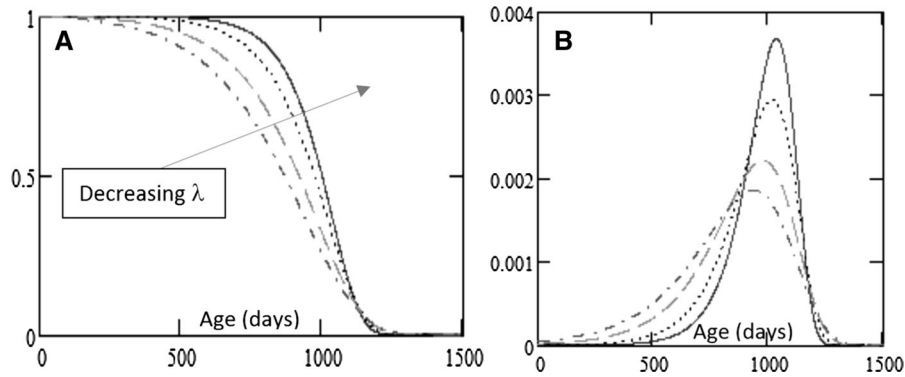


Fig. 8 Changes in the survival curves (a) and lifespan distributions (b), which conform to GM supplemented with a CEM. With decreasing λ , CDFs become increasingly “rectangular”. Note the similarity of this effect with the effect of SMC,

i.e., an apparent correlation between γ and λ , which emerges when data that conform to GMM are treated according to GM (Fig. 5)

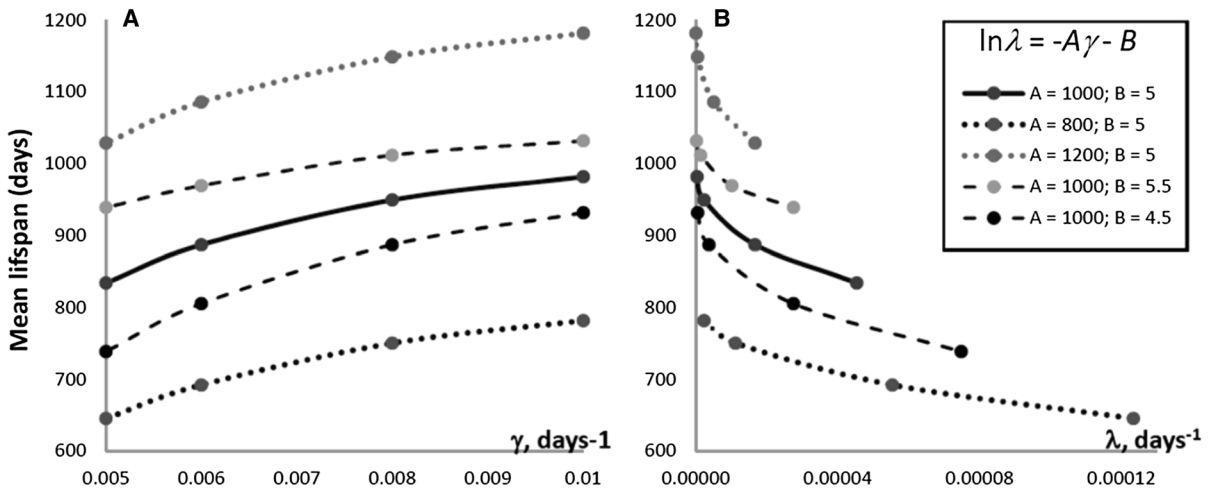


Fig. 9 The plots of mean lifespan versus γ (a) or (b) upon the condition that, because of the compensation effect of mortality, GM parameters are correlated as shown in the box

formally lifespan in such a population might be increased by accelerating the rate of aging, this contradicts common sense too overtly to be considered as the primary cause of the observed increase in longevity. In physiological terms, an acceleration of aging may be associated with an increase in the mean lifespan only if the acceleration is caused by the relocation of body resources to withstanding the causes of death away from mitigating body aging, that is, the acceleration of aging is the by-result of increasing the initial viability.

To make an idea about the contribution, if any, of a real CEM to apparent changes in γ and λ found in an

experiment of demographic analysis, it is reasonable to compare these changes with a SMC found in a series of historical controls or control subgroups obtained by randomizing a single control group. Such SMC will be entirely accounted for by the artifacts that must be eliminated from the apparent relationships between γ and λ that are observed upon an intervention into lifespan. Illustrative cases are presented in Figs. 10 and 11.

Figure 10 shows that the mitochondria-specific antioxidant SkQ1, which produces no significant effect according to log-rank test applied to the original Kaplan–Meier plots (Anisimov et al. 2011b), shifts the

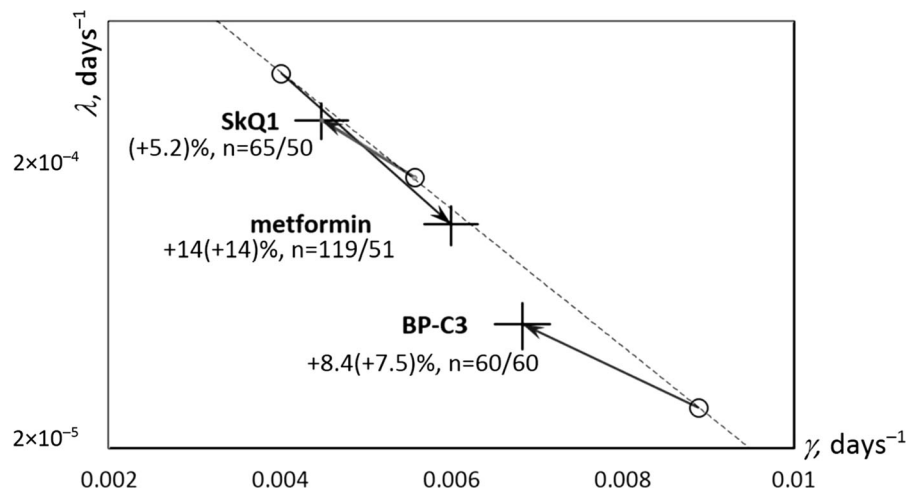


Fig. 10 The effects of metformin (Anisimov et al. 2011a), the antioxidant SkQ1 (Anisimov et al. 2011b), and the benzene polycarboxylic acid preparation BP-C3 (Anisimov et al. 2016) on aging rate γ and initial mortality λ in outbred female SHR mice. Controls data are shown with open circles. Effects are shown with arrows pointing at the respective 95% confidence

intervals (crosses) for resulting γ and λ . SMC (dashed line) is the same as in Fig. 7. Numbers at the names of drugs indicate percent changes in the mean lifespan, either reported (out of brackets), or calculated based on GM PDF (within brackets), and the numbers of control/experimental animals

respective experimental point away from the control point almost exactly along the SMC line.

The benzene polycarboxylic acid preparation BP-C3 (Anisimov et al. 2016) increases lifespan apparently by increasing λ and decreasing γ (Fig. 10). If CEM were at work, this combination would be impossible, as follows from Fig. 9. Indeed, since the increase in λ is less than should be expected based on SMC, the drug probably produces both of the possible beneficial effects: it decreases γ and at the same time decreases λ to an extent not counterbalanced by its artifactual increase. The net effect on survival is significant according the log-rank test used in the source publication. In Fig. 10, the significance of the effects is manifested in that SMC line is outside of 95% confidence limits of the coordinates of the experimental point.

The effect of BP-C3 is significant while being smaller than the effect of metformin, whose significance is marginal if judged by the deviation of the experimental point from the SMC line. Metformin apparently decreases λ and increases γ in SHR mice. Because the shift of the experimental point from the control is close to SMC line, it is hard to define based exclusively on statistics what really happens; therefore, the biological plausibility should be considered. In the original publication, log-rank test applied to

survival data showed that differences between control and experiment are statistically insignificant. In Fig. 10, SMC line crosses the borders of 95% confidence limits of the coordinates of the experimental point. At difference from BP-C3, which in mechanistic terms may supplement the antioxidant defenses, metformin by its effect on cell signaling pathways (Wu et al. 2017) is likely to shift the allocation of body resources in favor of ongoing viability versus protection from its deterioration (aging), which is consistent with CEM. If it were really so, the experimental point would be significantly below the SMC line. However, the shift occurred almost strictly along it. Therefore, metformin is likely to decrease λ , whereas the possible associated increase in γ is far less significant than it could appear if no account for the artificial components of SMC were taken. To conclude, there is no evidence that CEM is at work in this case.

The effect of metformin on γ and λ according to data obtained with female 129/Sv mice (Anisimov et al. 2010b) is shown in Fig. 11. An apparent decrease in λ is associated with an increase in γ . Both changes shift the experimental point along the SMC line, which is consistent with the lack of a significant effect on survival according to the log-rank test applied to data in the original publication.

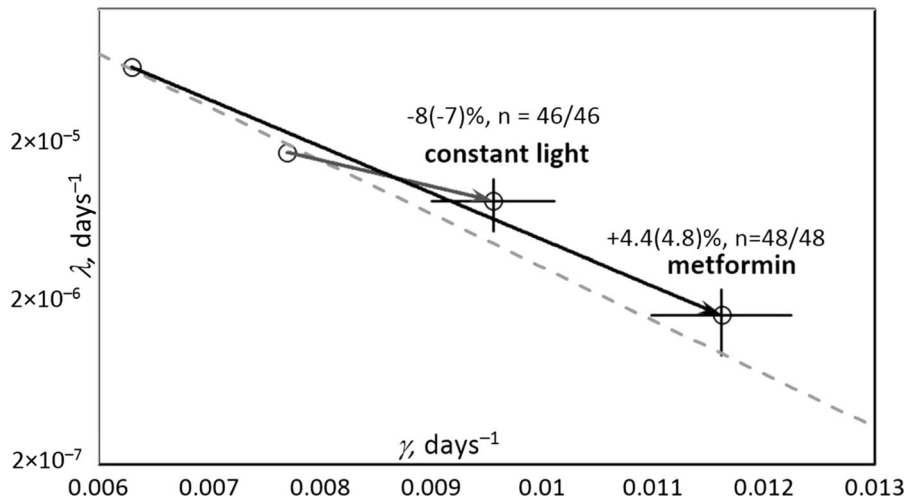


Fig. 11 The effects of metformin (Anisimov et al. 2010b) and constant illumination (Popovich et al. 2013) on γ and λ in 129/Sv mice with accounts of the same SM correlation (dashed line) as shown in Fig. 7. Captions are the same as in Fig. 10

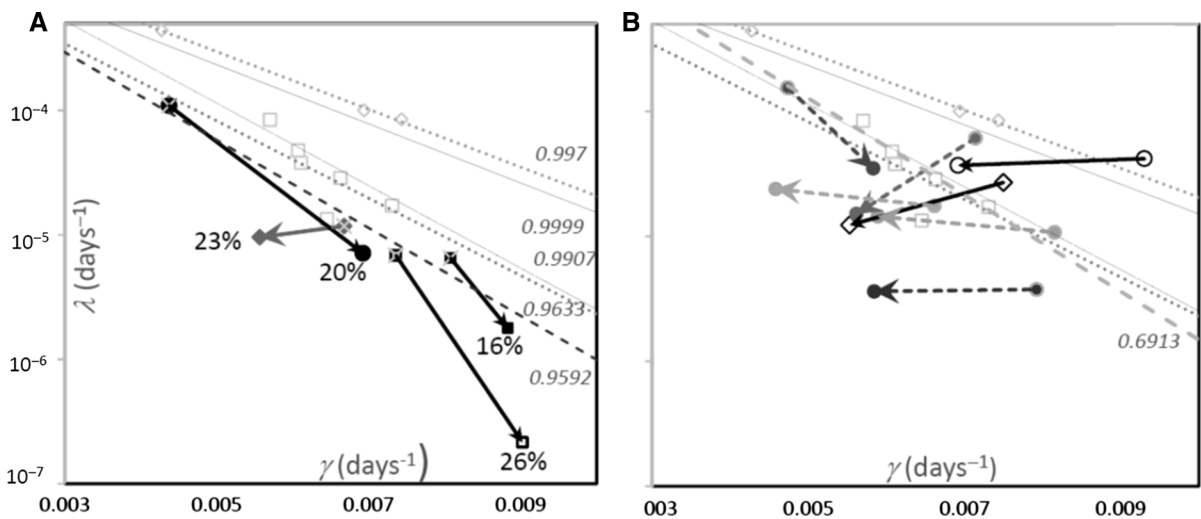


Fig. 12 Changes in GM parameters derived from published experiments where mTOR functions were attenuated (a) or calorie restriction was used (b) to extend lifespan in female mice. Thin lines show SMC obtained for 129/Sv and SHR mice bred at the author’s laboratory. Dotted lines show SMC based on datasets related to mice bred at Jackson Laboratories. Dashed lines shows SMC derived from all control published survival curves in a or from control survival curves published in (Turturo et al. 1999) in b. Italics show determination coefficients r^2 for respective SMC regressions. Percent numbers

show increases in the median lifespan. Symbols in a: Circles—mTOR gene mutation (Wu et al. 2013); open squares—rapamycin administration (Miller et al. 2014); filled squares—rapamycin administration (Fok et al. 2014); diamonds—rapamycin and metformin administration (Strong et al. 2016); arrows are directed from control to experimental points. Symbols in b: Dashed arrows are directed from control to experimental points derived from (Turturo et al. 1999), and solid arrows relate to data derived from (Mitchell et al. 2016). See the main text for additional comments

The effect of constant illumination observed in 129/Sv mice (Popovich et al. 2013) includes an increase in γ and a decrease in λ (Fig. 11), the latter, however, being significantly smaller than what may be

attributed to the artifactual component of SMC. Therefore, constant illumination results in an actual increase in λ , which is not counterbalanced by artifacts, and thus is likely to decrease lifespan by

increasing both, the initial mortality and the rate of aging. The effect was estimated as significant based on the log-rank test in the source publication, and it is expressed as a significant deviation of the experimental point from the SMC line in Fig. 11. There is no CEM at work in the present case, too.

The plots shown in Figs. 10 and 11 prompted us to check whether the approach used to construct them is applicable to published data on lifespan interventions in mice. Figure 12 shows the estimates of γ and λ extracted from published survival curves obtained in experiments where lifespan in mice was increased by mTOR gene mutation (Wu et al. 2013), lifelong rapamycin administration alone (Fok et al. 2014; Miller et al. 2014) or supplemented with metformin (Strong et al. 2016), and calorie restriction (CR) (Mitchell et al. 2016; Turturro et al. 1999). Because all our data relate to female mice, published data on female mice only were used. For any of the murine strains used to construct Fig. 12, no data are available about a SMC, which is constructed using data related to several control samples studied under constant conditions. However, because the two SMCs based on data about mice bred at Jackson Laboratories closely match the SMC based on our data about 128/Sv and SHR mice, it is plausible that these SMCs may be used to check against them the changes in γ and λ related to other murine strains. Remarkably, SMC based on the published datasets (dashed and dotted lines in Fig. 12) are almost parallel with SMC related to 129/Sv and SHR mice (thin pale lines).

All interventions used to attenuate the functions of mTOR (Fig. 12a) apparently reduce λ . Three of them are associated with apparent increases in γ . However, since the increases do not exceed what might be expected based on the lines that are parallel to SMC and cross the control points, the actual effects are decreases in γ , which are significant enough to be not obscured by SMC. Therefore, there is no CEM in mTOR-attenuating experiments. Each time both, γ and λ , actually decrease, contrarily to CEM.

If CEM were at work in experiments where CR increased lifespan in mice, no increases in λ would be possible. In the cases of apparent decreases in λ they are associated with decreases in γ (Fig. 12b), which again is inconsistent with CEM. In the only case where a decrease in λ is associated with an increase in γ , the latter is smaller than the artificial component of SMC would produce; therefore, γ actually decreases in this

case. Altogether, this suggests that both, λ and γ , decrease upon CR, contrarily to CEM.

Notably, the patterns of vectors in Fig. 12a and b are different. The dominant trend in B is leftward, whereas in A it is descending rightward, similar to the effect of metformin administration (Figs. 10, 11). Thus, the effects of CR in female mice are different from the effects of drugs often labeled as CR mimetics, as e.g. in (Calvert et al. 2016; Lee and Min 2013). It follows from comparing the two panels of Fig. 12 that increases in lifespan upon CR are mainly caused by decelerated aging, whereas upon mTOR attenuation, by decreased initial mortality. In either case, no manifestations of CEM are evident.

General discussion

“The term ‘effective thinking’ is used to refer to the philosophy of placing emphasis on the interpretation of overall effect size in terms of biological importance rather than statistical significance” (Nakagawa et al. 2017). With regard to assessing the overall effect size in terms of changes in the rate of aging in mice, the first conclusion from our analysis is that it is important each time to be sure that aging indeed contributes to the observed pattern of survival. It clearly does not in the case of HER2/neu mice, and its contribution is dubious in the case of SAMP mice.

In control SHR and 129/Sv mice, the negative correlations between the terms of the Gompertz model (GM), which are interpretable as the demographic rate of aging and the inverse of the mean initial vitality ($\ln \lambda = -A\gamma - B$), appears to be an artifact. This observation may seem merely a confirmation of a common, although not generally recognized, catch in the way of approximating any noisy data with any nonlinear multi-parametric model (Johnson 2000). The specific form of the artifact, which is known as the Strehler-Mildvan correlation (SMC) emerges because of the pre- and post-exponent positions of GM terms (Tarkhov et al. 2017). Even if only an artifact, it is important not to dismiss it when survival or mortality data are used to discriminate changes in the initial physiological viability (which in populations is reflected in an inverse manner in the pre-exponent term λ of GM) and the rate of its age-associated decline (which is captured by γ). Indeed, it has been suggested that SMC, irrespectively of its source,

should be accounted for in making inferences about aging in extinct human populations (Sasaki and Kondo 2016).

The significance of SMC could be limited to the above technical matters if not for the fact that it had been predicted based on physical considerations by B. Strehler and A. Mildvan before it was sought for and found by them in human mortality data (Strehler and Mildvan 1960). That is, a theoretical prediction was confirmed rather than a post hoc explanation was suggested for a serendipitous observation, which eventually turned out to be artifactual. In fact, the initially reported SMC found by applying GM to human data is almost fully attributable to ignoring changes in the Makeham term C of GMM. However, the persistence of reciprocity of changes in γ and λ even if every effort is undertaken to minimize the contribution of changes in C to human mortality patterns, as e.g. in (Golubev 2009), justifies a more elaborate discussion.

It is plausible biologically that when more of limited resources is allocated to protection from death, less is left to repair damage that can accumulate in the means of protection. This tradeoff is generally consistent with the disposable soma theory of aging (Kirkwood 1977; Drenos and Kirkwood 2005), although its main emphasis is on tradeoffs between self-repair and self-reproduction rather than self-protection. Further, the antagonistic pleiotropy theory of the evolutionary origins of aging (Williams 1957) implies that an improvement in the early fitness may be supported by natural selection even if the improvement is associated with a greater decline in the later fitness, such as because of accelerated senescence (Williams 1957; Rose 1985; Golubev et al. 2017b). It is also plausible that senescence-causing damage may be a byproduct of the execution of self-protecting functions and thus may be potentiated by increasing the power applied to self-protection. This conforms to the rate-of-living theory [see (Hulbert et al. 2007)]. Therefore, as far as, in populations, the initial viability is manifested in a reverse manner in the initial mortality, and aging, in the age-dependent increase in mortality, the respective demographic parameters must be correlated. Thus, the three basic theories of aging are generally consistent with SMC (Golubev et al. 2017a, b).

On the other hand, it has been shown in (Burger and Missov 2016; Tarkhov et al. 2017) and confirmed here

with murine survival data (see Figs. 6, 7 and their discussion) that SMC can emerge as a byproduct of the mathematical treatment of survival data. A challenge therefore is to distinguish in SMC an artifact, which may be unidirectional with a real phenomenon, from the phenomenon as it is. The real phenomenon comprises tradeoffs between the physiological parameters responsible for the capability of self-protection and for protecting this capability from age-dependent deterioration. The compensation effects of mortality (CEM) may be thought of as showing how these physiological tradeoffs are manifested in demographic parameters. The first step in distinguishing of CEM within SMC was the recognition (Gavrilov and Gavrilova 1991; Golubev 2004) that, when survival data in a population where the age-independent (background, extrinsic) mortality (which is captured by the term C of GMM) is significant and variable are treated as if the data conform to the pure GM, (where there is no C at all), the reciprocal changes in the demographic parameters λ and γ thus found may be mistaken for the manifestation of a real correlation between the initial physiological viability and the rate of its decline, i.e., the physiological aging. This situation is illustrated here in Fig. 6.

Although the need to distinguish CEM in an observed SMC is explicitly proclaimed by some authors (Golubev 2009; Strulik and Vollmer 2013), the lack of consensus on this issue seems to be responsible for much of current confusion around relationships between changes in GM and GMM parameters.

There is also no consensus concerning an even more fundamental issue: whether GM and/or GMM is a manifestation of some natural laws behind it, or it is merely a tool adopted by convention for treating survival and mortality data. The latter attitude persists through decades as it is evident from the two claims: "... the lack of simple and efficient procedure to comparing different mortality model forced the use of the Gompertz model... which may not apply to the majority of experimental systems" (Pletcher 1999) and "We based our simulations on the Gompertz equation because of its extensive use in the past and because it models death times reasonably well across many species. ...Insofar as the Gompertz equation is not based on a mathematical formulation of how aging works, its parameters ... do not necessarily describe

any biological or molecular entity” (Petrascheck and Miller 2017).

Our numerical experiments were motivated by the belief that GMM is based on a mathematical formulation of how aging works in parent populations represented by experimental samples. It has been argued that the constant term of GMM is not just an additional parameter needed for nothing more than improving model fits to data, but an integral part of a generalized Gompertz-Makeham law (GGML) (Golubev 2009), which is not limited by the assumption that the age-dependent physiological decline is linear (that is, aging rate is constant). GGML is not a consequence of what is defined in (Hamilton 1966) as “molding of senescence by natural selection”, but rather a manifestation of fundamental constraints imposed on natural selection by the physicochemical properties of the material to be “molded” (Golubev 2009; Golubev et al. 2017a,b). The canonical GM is but an idealized approximation of GGML. Therefore, the real survival and mortality patterns just never can fully conform to GM, GMM, LGM or any combination thereof, and this unavoidably will make the estimates of their parameters correlated, irrespective of the noisiness of data and the accuracy of accounting for any biases introduced by mathematical treatment (sampling, binning etc.). This situation is exemplified here in Fig. 6. Another example of the same may be found in the results of extraction of GMM parameters from data on human mortality (Golubev 2012): subtle details in the trends of changes in γ and $\ln\lambda$ mirror each other too accurately to be real. These apparent closely correlated local minimums in $\ln\lambda$ versus time plots and simultaneous local maximums in γ versus time plots were suggested to result from changes in the conditions of human living, which alter the properties of successive overlapping generations. Indeed, survival and mortality patterns strictly conforming to GMM may be observed in period data only if the properties of all generations that comprise a population under study are identical, and in a longitudinal study, only if conditions over the whole observation period are constant. The question remains, however, whether during the last century there was (and still is now) a general trend of increasing initial human vitality associated with increasing human aging rate, or, on the contrary, there was a decrease in the rate of human aging.

The issues therefore are not only how much SMC is an artifact and whether CEM may contribute to SMC, but also what are the conditions for physiology behind CEM to work.

Under the conditions of speciation by natural selection, both γ and λ may decrease in parallel, otherwise no long-lived species could have evolved. For example, in a series of less to more advanced primates—from baboons through apes to humans—both λ and γ decrease in a manner consistent with a linear positive correlation (Bronikowski et al. 2011) where there is nothing of CEM. Upon experimental evolution of aging and longevity in *D. melanogaster* and *C. elegans*, increases in longevity brought about by decreases in the rate of aging are found associated with compromised performance not always but often enough (Briga and Verhulst 2015; Wit et al. 2013; Zwaan et al. 1995) to disregard not this manifestation of CEM. Therefore, a closer reexamination of these results, including making estimates of GM parameters with account for possible artifactual component of SMC is warranted. The observations that under stressful conditions long-lived mutants are uncompetitive versus their wild-type counterparts (Briga and Verhulst 2015) generally conform to CEM. Comparing different strains of laboratory mice, many of which result from unintentional experimental evolution based on selection for high and early-onset fertility, does reveal SMC (e.g., Figs. 7, 12); however, it is still not clear does a true CEM contribute to it.

Of special interest in this regard are changes in GM parameters upon direct experimental interventions in the lifespans of mammals, which are the most relevant to developing approaches applicable to humans and, at the same time, are the most vulnerable to all sorts of artifacts associated with relatively small experimental samples and magnitudes of effects (Eakin et al. 1995, Bokov et al. 2017; Petrascheck and Miller 2017). Figures 10, 11, and 12 are instructive in this regard by showing that the apparent reciprocity of changes in GM parameters may be reduced to none and that the apparent changes in lifespan may be associated with unidirectional changes in GM parameters, provided the artifactual component of SMC is taken into account. The lack of due attention to this aspect of data analysis in experimental aging research is a likely source of current controversies around the contribution of changes in the rate of aging to changes in lifespan.

As concerns the still unsettled disputes, e.g., around the effects of calorie restriction in mice, inferences from Fig. 12b corroborate the conclusion that the main effect of CR in female mice is to slow down aging (Simons et al. 2013), contrary to the claim that both γ and λ are reduced by CR (Nakagawa et al. 2012). In this regard, CR differs from mTOR attenuation, which in female mice decreases λ , as may be inferred from Fig. 12a. However, this decrease is not associated with increasing γ , that is, there is no CEM, contrary to the conclusion expressed in (Shen et al. 2017) in terms of SMC without differentiating its artifactual and substantive components. On the whole, our analysis is consistent with the conclusion about differences between mTOR attenuation and CR in mice, which was drawn in (Garratt et al. 2016) based on a meta-analysis of published experiments. This does not rule out the possibility that it may be otherwise in male mice, not to mention fruit flies and nematodes. Checking all relevant findings against respective artificial components of SMC is beyond the scope of the present study, which is limited to female mice.

Our final comments are on the relevance of distinguishing CEM within SMC to human survival patterns. The prolongation of healthy human lifespan, which is the ultimate objective for gerontology, is conventionally described in terms of “rectangularized” survival curves or “compressed mortality” patterns (Cheung and Robine 2007; Wilmoth and Horiuchi 1999). In terms of GMM, “rectangularization” may be achieved by a decrease in the age-independent (background) mortality and/or by an increase in the initial physiological vitality (which is manifested as a decrease in the initial demographic mortality), but only in association with an increase in the rate of physiological aging (which is manifested as an accelerated increase in the age dependent mortality). The latter option is nothing else but CEM (compare Figs. 5, 8). The two options are not easily distinguishable based on demographic data. C clearly was decreasing in human populations over the last 100 years when unprecedented increases in human life expectancy took place (Golubev 2012). It is far less clear whether there was an increase in the initial human robustness, which in human populations is reflected in decreasing λ , and whether this occurred at the expense of increasing γ . A smaller λ is obviously a desirable outcome of interventions intended to improve human conditions through the whole lifespan

and thus to increase it in a way, which is consistent with the rectangularization of survival curves, that is with CEM. At the same time, a smaller γ will make a survival curve less “rectangular” and mortality less “compressed”.

Altogether, this makes it doubtful whether aging is what should be targeted in order to increase human health and life spans. The term “anti-aging” thus appears essentially misleading (even if commercially attractive) with regard to translating advances in experimental gerontology into practice. However, the rate of aging does appear to decrease when the artifactual components of SMC are accounted for, even in experiments where aging seems to be accelerated if the components are not taken into account (Figs. 10, 11, and 12).

Therefore, to ascertain what really happens upon experimental interventions in lifespan, it is reasonable to check apparent changes in GM parameters against a predefined SMC based on a series of historical controls, or on randomizing a sufficiently big control group into several subgroups.

In a recent paper (Petrascheck and Miller 2017), quantitative relationships between the numbers of animals in experimental groups and the ability to detect real changes in longevity have been established based on GM (which however was treated as merely a handy tool). For example, a real increase within 16% can be detected in not more than a half of experiments using 100 animals in each of two groups, control and experimental. To improve “the power of detection” of real effects, the sizes of experimental groups must be increased.

With more than 100 control animals, it is possible to randomize them into four or more reasonably large subgroups and construct the respective artifactual component of Strehler-Mildvan correlation for checking against it whether actual changes in the rate of aging or the initial vitality are responsible for the observed changes in the mean lifespan, and whether a correlation, if any, between the two contributors to changes in lifespan is consistent with the compensation effect of mortality.

Acknowledgements The authors would like to thank Drs. Maria Yurova and Margarita Tyndyk (N.N. Petrov Research Institute of Oncology) and Irina Popovich (Institute of Bioregulation and Gerontology, Saint Petersburg, Russia) for commenting their data. This publication is partly supported by

the Grant RNF No 16-15-00142 from Russian Science Foundation.

Compliance with ethical standards

Conflicts of interest The authors declare no conflicts of interest.

References

- Anderson JJ, Li T, Sharrow DJ (2017) Insights into mortality patterns and causes of death through a process point of view model. *Biogerontology* 18:149–170. <https://doi.org/10.1007/s10522-016-9669-1>
- Anisimov VN, Alimova IN, Baturin DA, Popovich IG, Zabezhinski MA, Manton KG, Semenchenko AV, Yashin AI (2003) The effect of melatonin treatment regimen on mammary adenocarcinoma development in HER-2/neu transgenic mice. *Int J Cancer* 103:300–305. <https://doi.org/10.1002/ijc.10827>
- Anisimov VN, Berstein LM, Egorin PA, Piskunova TS, Popovich IG, Zabezhinski MA, Kovalenko IG, Poroshina TE, Semenchenko AV, Provinciali M, Re F, Franceschi C (2005) Effect of metformin on life span and on the development of spontaneous mammary tumors in HER-2/neu transgenic mice. *Exper Gerontol* 40:685–693. <https://doi.org/10.1016/j.exger.2005.07.007>
- Anisimov VN, Egorin PA, Piskunova TS, Popovich IG, Tyndyk ML, Yurova MN, Zabezhinski MA, Anikin IV, Karkach AS, Romanyukha AA (2010a) Metformin extends life span of HER-2/neu transgenic mice and in combination with melatonin inhibits growth of transplantable tumors in vivo. *Cell Cycle* 9:188–197. <https://doi.org/10.4161/cc.9.1.10407>
- Anisimov VN, Piskunova TS, Popovich IG, Zabezhinski MA, Tyndyk ML, Egorin PA, Yurova MN, Rosenfeld SV, Semenchenko AV, Kovalenko IG, Poroshina TE, Berstein LM (2010b) Gender differences in metformin effect on aging, life span and spontaneous tumorigenesis in 129/Sv mice. *Aging* 2:945–958
- Anisimov VN, Zabezhinski MA, Popovich IG, Piskunova TS, Semenchenko AV, Tyndyk ML, Yurova MN, Antoch MP, Blagosklonny MV (2010c) Rapamycin extends maximal lifespan in cancer-prone mice. *Am J Pathol* 176:2092–2097. <https://doi.org/10.2353/ajpath.2010.091050>
- Anisimov VN, Berstein LM, Popovich IG, Zabezhinski MA, Egorin PA, Piskunova TS, Semenchenko AV, Tyndyk ML, Yurova MN, Kovalenko IG, Poroshina TE (2011a) If started early in life, metformin treatment increases life span and postpones tumors in female SHR mice. *Aging* 3:148–157
- Anisimov VN, Egorov MV, Krasilshchikova MS, Lyamzaev KG, Manskikh VN, Moshkin MP, Novikov EA, Popovich IG, Rogovin KA, Shabalina IG, Shekarova ON, Skulachev MV, Titova TV, Vygodin VA, Vysokikh MY, Yurova MN, Zabezhinsky MA, Skulachev VP (2011b) Effects of the mitochondria-targeted antioxidant SkQ1 on lifespan of rodents. *Aging* 3:1110–1119. <https://doi.org/10.18632/aging.100404>
- Anisimov VN, Popovich IG, Zabezhinski MA (2013) Methods of testing pharmacological drugs effects on aging and life span in mice. *Methods Mol Biol* (Clifton, NJ) 1048:145–160. https://doi.org/10.1007/978-1-62703-556-9_12
- Anisimov VN, Popovich IG, Zabezhinski MA, Egorin PA, Yurova MN, Semenchenko AV, Tyndyk ML, Panchenko AV, Trashkov AP, Vasiliev AG, Khaitsev NV (2015) Sex differences in aging, life span and spontaneous tumorigenesis in 129/Sv mice neonatally exposed to metformin. *Cell Cycle* 14:46–55. <https://doi.org/10.4161/15384101.2014.973308>
- Anisimov VN, Popovich IG, Zabezhinski MA, Yurova MN, Tyndyk ML, Anikin IV, Egorin PA, Baldueva IA, Fedoros EI, Pigarev SE, Panchenko AV (2016) Polyphenolic drug composition based on benzenepolycarboxylic acids (BP-C3) increases life span and inhibits spontaneous tumorigenesis in female SHR mice. *Aging* 8:1866–1875. <https://doi.org/10.18632/aging.101024>
- Avraam D, Arnold S, Vasieva O, Vasiev B (2016) On the heterogeneity of human populations as reflected by mortality dynamics. *Aging* 8:3045–3064. <https://doi.org/10.18632/aging.101112>
- Bokov AF, Manuel LS, Tirado-Ramos A, Gelfond JA, Pletcher SD (2017) Biologically relevant simulations for validating risk models under small-sample conditions. In: 2017 IEEE symposium on computers and communications (ISCC), 3–6 July 2017, pp 290–295. <https://doi.org/10.1109/iscc.2017.8024544>
- Brayton CF, Treuting PM, Ward JM (2012) Pathobiology of aging mice and GEM: background strains and experimental design. *Vet Pathol* 49:85–105. <https://doi.org/10.1177/0300985811430696>
- Breslow NE (1992) Introduction to Kaplan and Meier (1958) Nonparametric estimation from incomplete observations. In: Kotz S, Johnson NL (eds) Breakthroughs in statistics: methodology and distribution. Springer New York, New York, pp 311–318. https://doi.org/10.1007/978-1-4612-4380-9_24
- Briga M, Verhulst S (2015) What can long-lived mutants tell us about mechanisms causing aging and lifespan variation in natural environments? *Exper Gerontol* 71:21–26. <https://doi.org/10.1016/j.exger.2015.09.002>
- Bronikowski AM, Altmann J, Brockman DK, Cords M, Fedigan LM, Pusey A, Stoinski T, Morris WF, Strier KB, Alberts SC (2011) Aging in the natural world: comparative data reveal similar mortality patterns across primates. *Science* 331:1325–1328. <https://doi.org/10.1126/science.1201571>
- Burger O, Missov TI (2016) Evolutionary theory of ageing and the problem of correlated Gompertz parameters. *J Theor Biol* 408:34–41. <https://doi.org/10.1016/j.jtbi.2016.08.002>
- Calvert S, Tacutu R, Sharifi S, Teixeira R, Ghosh P, de Magalhaes JP (2016) A network pharmacology approach reveals new candidate caloric restriction mimetics in *C. elegans*. *Aging Cell* 15:256–266. <https://doi.org/10.1111/acer.12432>
- Carter TA, Greenhall JA, Yoshida S, Fuchs S, Helton R, Swaroop A, Lockhart DJ, Barlow C (2005) Mechanisms of

- aging in senescence-accelerated mice. *Genome Biol* 6:R48. <https://doi.org/10.1186/gb-2005-6-6-r48>
- Chapman JW, O'Callaghan CJ, Hu N, Ding K, Yothers GA, Catalano PJ, Shi Q, Gray RG, O'Connell MJ, Sargent DJ (2013) Innovative estimation of survival using log-normal survival modelling on ACCENT database *Brit J. Cancer* 108:784–790. <https://doi.org/10.1038/bjc.2013.34>
- Cheung SL, Robine JM (2007) Increase in common longevity and the compression of mortality: the case of Japan. *Popul Stud (Camb)* 61:85–97. <https://doi.org/10.1080/00324720601103833>
- de Beer J, Janssen F (2016) A new parametric model to assess delay and compression of mortality. *Population Health Metrics* 14:46. <https://doi.org/10.1186/s12963-016-0113-1>
- de Magalhaes JP, Cabral JA, Magalhaes D (2005) The influence of genes on the aging process of mice: a statistical assessment of the genetics of aging. *Genetics* 169:265–274. <https://doi.org/10.1534/genetics.104.032292>
- Dixon WJ, Mood AM (1946) The statistical sign test. *J Am Stat Assoc* 41:557–566. <https://doi.org/10.2307/2280577>
- Drenos F, Kirkwood TB (2005) Modelling the disposable soma theory of ageing. *Mech Ageing Dev* 126:99–103. <https://doi.org/10.1016/j.mad.2004.09.026>
- Eakin T, Shouman R, Qi Y, Liu G, Witten M (1995) Estimating parametric survival model parameters in gerontological aging studies: methodological problems and insights. *J Gerontol Ser A Biol Sci Med Sci* 50:B166–176
- Finkelstein M (2012) Discussing the Strehler-Mildvan model of mortality. *Demogr Res* 26:1
- Fok WC, Chen Y, Bokov A, Zhang Y, Salmon AB, Diaz V, Javors M, Wood WH 3rd, Zhang Y, Becker KG, Pérez VI, Richardson A (2014) Mice fed rapamycin have an increase in lifespan associated with major changes in the liver transcriptome *PLOS ONE* 9:e83988. <https://doi.org/10.1371/journal.pone.0083988>
- Garratt M, Nakagawa S, Simons MJ (2016) Comparative idiosyncrasies in life extension by reduced mTOR signalling and its distinctiveness from dietary restriction. *Aging Cell* 15:737–743. <https://doi.org/10.1111/accel.12489>
- Gavrilov LA, Gavrilova NS (1991) *The biology of life span: a quantitative approach*. Harwood Academic Publisher, New York
- Golubev A (2004) Does Makeham make sense? *Biogerontology* 5:159–167
- Golubev A (2009) How could the Gompertz-Makeham law evolve. *J Theor Biol* 258:1–17. <https://doi.org/10.1016/j.jtbi.2009.01.009>
- Golubev AG (2012) The issue of the feasibility of a general theory of aging. III. Theory and practice of aging *Adv Gerontol* 2:109–119. <https://doi.org/10.1134/s207905701206001x>
- Golubev A (2016) Applications and implications of the exponentially modified gamma distribution as a model for time variabilities related to cell proliferation and gene expression. *J Theor Biol* 393:203–217. <https://doi.org/10.1016/j.jtbi.2015.12.027>
- Golubev A, Hanson AD, Gladyshev VN (2017a) Non-enzymatic molecular damage as a prototypic driver of aging. *J Biol Chem* 292:6029–6038. <https://doi.org/10.1074/jbc.R116.751164>
- Golubev A, Hanson AD, Gladyshev VN (2017b) A tale of two concepts: harmonizing the free radical and antagonistic pleiotropy theories of aging. *Antioxid Redox Signal*. <https://doi.org/10.1089/ars.2017.7105>
- Hamilton WD (1966) The moulding of senescence by natural selection. *J Theor Biol* 12:12–45. [https://doi.org/10.1016/0022-5193\(66\)90184-6](https://doi.org/10.1016/0022-5193(66)90184-6)
- Hughes BG, Hekimi S (2016) Different mechanisms of longevity in long-lived mouse and *Caenorhabditis elegans* mutants revealed by statistical analysis of mortality rates. *Genetics* 204:905–920. <https://doi.org/10.1534/genetics.116.192369>
- Hulbert AJ, Pamplona R, Buffenstein R, Buttemer WA (2007) Life and death: metabolic rate, membrane composition, and life span of animals. *Physiol Rev* 87:1175–1213. <https://doi.org/10.1152/physrev.00047.2006>
- Iurova MH, Zabezhinskii MA, Piskunova TS, Tyndyk ML, Popovich IG, anisimov VN (2010) [The effect of mitochondria targeted antioxidant SkQ1 on aging, life span and spontaneous carcinogenesis in three mice strains]. *Adv Gerontol = Uspekhi gerontologii* 23:430–441
- Johnson ML (2000) Parameter correlations while curve fitting *Meth Enzymol* 321:424–446
- Kirkwood TB (1977) Evolution of ageing *Nature* 270:301–304
- Kirkwood TBL (2015) Deciphering death: a commentary on Gompertz (1825) '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'. *Philos Trans R Soc B: Biol Sci* 370:1666. <https://doi.org/10.1098/rstb.2014.0379>
- Lee S-H, Min K-J (2013) Caloric restriction and its mimetics *BMB Reports* 46:181–187. <https://doi.org/10.5483/BMBRep.2013.46.4.033>
- Liang Y, Liu C, Lu M, Dong Q, Wang Z, Wang Z, Xiong W, Zhang N, Zhou J, Liu Q, Wang X, Wang Z (2018) Calorie restriction is the most reasonable anti-ageing intervention: a meta-analysis of survival curves *Sci Rep* 8:5779. <https://doi.org/10.1038/s41598-018-24146-z>
- Limpert E, Stahel WA (2011) Problems with using the normal distribution—and ways to improve quality and efficiency of data analysis. *PLoS ONE* 6:e21403. <https://doi.org/10.1371/journal.pone.0021403>
- Limpert E, Stahel WA (2017) The log-normal distribution *Significance* 14:8–9. <https://doi.org/10.1111/j.1740-9713.2017.00993.x>
- Miller RA, Harrison DE, Astle CM, Fernandez E, Flurkey K, Han M, Javors MA, Li X, Nadon NL, Nelson JF, Pletcher S, Salmon AB, Sharp ZD, Van Roekel S, Winkelman L, Strong R (2014) Rapamycin-mediated lifespan increase in mice is dose and sex dependent and metabolically distinct from dietary restriction. *Aging Cell* 13:468–477. <https://doi.org/10.1111/accel.12194>
- Mitchell SJ, Madrigal-Matute J, Scheibye-Knudsen M, Fang E, Aon M, González-Reyes JA, Cortassa S, Kaushik S, Gonzalez-Freire M, Patel B, Wahl D, Ali A, Calvo-Rubio M, Burón MI, Guterrez V, Ward TM, Palacios HH, Cai H, Frederick DW, Hine C, Broeskamp F, Habering L, Dawson J, Beasley TM, Wan J, Ikeno Y, Hubbard G, Becker KG, Zhang Y, Bohr VA, Longo DL, Navas P, Ferrucci L, Sinclair DA, Cohen P, Egan JM, Mitchell JR, Baur JA, Allison DB, Anson RM, Villalba JM, Madeo F, Cuervo AM,

- Pearson KJ, Ingram DK, Bernier M, de Cabo R (2016) Effects of sex, strain, and energy intake on hallmarks of aging in mice. *Cell Metab* 23:1093–1112. <https://doi.org/10.1016/j.cmet.2016.05.027>
- Nakagawa S, Lagisz M, Hector KL, Spencer HG (2012) Comparative and meta-analytic insights into life extension via dietary restriction. *Aging Cell* 11:401–409. <https://doi.org/10.1111/j.1474-9726.2012.00798.x>
- Nakagawa S, Noble DWA, Senior AM, Lagisz M (2017) Meta-evaluation of meta-analysis: ten appraisal questions for biologists. *BMC Biol* 15:18. <https://doi.org/10.1186/s12915-017-0357-7>
- Olshansky SJ, Carnes BA (1997) Ever since Gompertz Demography 34:1–15
- Panchenko AV, Popovich IG, Trashkov AP, Egormin PA, Yurova MN, Tyndyk ML, Gubareva EA, Artyukin IN, Vasiliev AG, Khaitsev NV, Zabezhinski MA, Anisimov VN (2016) Biomarkers of aging, life span and spontaneous carcinogenesis in the wild type and HER-2 transgenic FVB/N female mice. *Biogerontology* 17:317–324. <https://doi.org/10.1007/s10522-015-9611-y>
- Petrasccheck M, Miller D (2017) Computational analysis of lifespan experiment reproducibility BioRxiv. <https://doi.org/10.1101/107417>
- Pletcher (1999) Model fitting and hypothesis testing for age-specific mortality data. *J Evol Biol* 12:430–439. <https://doi.org/10.1046/j.1420-9101.1999.00058.x>
- Pletcher SD, Khazaali AA, Curtsinger JW (2000) Why do life spans differ? Partitioning mean longevity differences in terms of age-specific mortality parameters. *J Gerontol Ser A Biol Sci Med Sci* 55:B381–389
- Popovich IG, Zabezhinskii MA, Egormin PA, Anikin IV, Tyndyk ML, Semenchenko AV, Spasov AA, Anisimov VN (2004) [Effect of antidiabetic drug diabenol on parameters of biological age, life span and tumor development in NMRI and HER-2/neu mice]. *Adv Gerontol = Uspekhi gerontologii* 15:80–90
- Popovich IG, Zabezhinski MA, Panchenko AV, Piskunova TS, Semenchenko AV, Tyndyk ML, Yurova MN, Anisimov VN (2013) Exposure to light at night accelerates aging and spontaneous uterine carcinogenesis in female 129/Sv mice. *Cell Cycle* 12:1785–1790. <https://doi.org/10.4161/cc.24879>
- Popovich IG, Anisimov VN, Zabezhinski MA, Semenchenko AV, Tyndyk ML, Yurova MN, Blagosklonny MV (2014) Lifespan extension and cancer prevention in HER-2/neu transgenic mice treated with low intermittent doses of rapamycin. *Cancer Biol Ther* 15:586–592. <https://doi.org/10.4161/cbt.28164>
- Porquet D, Casadesus G, Bayod S, Vicente A, Canudas AM, Vilaplana J, Pelegri C, Sanfeliu C, Camins A, Pallas M, del Valle J (2013) Dietary resveratrol prevents Alzheimer's markers and increases life span in SAMP8. *Age (Dordrecht, Netherlands)* 35:1851–1865. <https://doi.org/10.1007/s11357-012-9489-4>
- Rose MR (1985) Life history evolution with antagonistic pleiotropy and overlapping generations *Theor Populat Biol* 28:342–358. [https://doi.org/10.1016/0040-5809\(85\)90034-6](https://doi.org/10.1016/0040-5809(85)90034-6)
- Royston P (2001) The lognormal distribution as a model for survival time in cancer, with an emphasis on prognostic factors. *Stat Neerl* 55:89–104. <https://doi.org/10.1111/1467-9574.00158>
- Sasaki T, Kondo O (2016) An informative prior probability distribution of the gompertz parameters for bayesian approaches in paleodemography. *Am J Phys Anthropol* 159:523–533. <https://doi.org/10.1002/ajpa.22891>
- Shen J, Landis GN, Tower J (2017) Multiple metazoan life-span interventions exhibit a sex-specific Strehler-Mildvan inverse relationship between initial mortality rate and age-dependent mortality rate acceleration. *J Gerontol Ser A Biol Sci Med Sci* 72:44–53. <https://doi.org/10.1093/gerona/glw005>
- Shouman R, Witten M (1995) Survival estimates and sample size: what can we conclude? *J Gerontol Ser A, Biol Sci Med Sci* 50:B177–185
- Simons MJ, Koch W, Verhulst S (2013) Dietary restriction of rodents decreases aging rate without affecting initial mortality rate—a meta-analysis. *Aging Cell* 12:410–414. <https://doi.org/10.1111/accel.12061>
- Spratt JS (1969) The lognormal frequency distribution and human cancer *J Surgicall Res* 9:151–157. [https://doi.org/10.1016/0022-4804\(69\)90046-8](https://doi.org/10.1016/0022-4804(69)90046-8)
- Strehler BL, Mildvan AS (1960) General theory of mortality and aging. *Science* 132:14–21
- Strong R, Miller RA, Antebi A, Astle CM, Bogue M, Denzel MS, Fernandez E, Flurkey K, Hamilton KL, Lamming DW, Javors MA, de Magalhães JP, Martinez PA, McCord JM, Miller BF, Müller M, Nelson JF, Ndukum J, Rainger GE, Richardson A, Sabatini DM, Salmon AB, Simpkins JW, Steegenga WT, Nadon NL, Harrison DE (2016) Longer lifespan in male mice treated with a weakly estrogenic agonist, an antioxidant, an α -glucosidase inhibitor or a Nrf2-inducer. *Aging Cell* 15:872–884. <https://doi.org/10.1111/accel.12496>
- Strulik H, Vollmer S (2013) Long-run trends of human aging and longevity *J Populat Econ* 26:1303–1323. <https://doi.org/10.1007/s00148-012-0459-z>
- Tabeau E (2001) A review of demographic forecasting models for mortality. In: Tabeau E, van den Berg Jeths A, Heathcote C (eds) *Forecasting mortality in developed countries: insights from a statistical, demographic and epidemiological perspective*. Springer Netherlands, Dordrecht, pp 1–32. https://doi.org/10.1007/0-306-47562-6_1
- Tai TH, Noymer A (2018) Models for estimating empirical Gompertz mortality: with an application to evolution of the Gompertzian slope population ecology e-published ahead of print. <https://doi.org/10.1007/s10144-018-0609-6>
- Takeda T, Hosokawa M, Takeshita S, Irino M, Higuchi K, Matsushita T, Tomita Y, Yasuhira K, Hamamoto H, Shimizu K, Ishii M, Yamamuro T (1981) A new murine model of accelerated senescence *Mech Ageing Develop* 17:183–194. [https://doi.org/10.1016/0047-6374\(81\)90084-1](https://doi.org/10.1016/0047-6374(81)90084-1)
- Tarkhov AE, Menshikov LI, Fedichev PO (2017) Strehler-Mildvan correlation is a degenerate manifold of Gompertz fit. *J Theor Biol* 416:180–189. <https://doi.org/10.1016/j.jtbi.2017.01.017>
- Turturro A, Witt WW, Lewis S, Hass BS, Lipman RD, Hart RW (1999) Growth curves and survival characteristics of the animals used in the biomarkers of aging program *J*

- Gerontol: Ser A 54:B492–B501. <https://doi.org/10.1093/gerona/54.11.B492>
- Tyndyk ML, Popovich IG, Anikin IV, Egormin PA, Iurova MN, Zabezhinskii MA, Anisimov VN (2012) The effect of SSH&H on the lifespan and spontaneous cancer development in transgenic mice with HER-2/neu mutation. *Vopr Onkol* 58:243–247
- Umezawa M, Takeda T, Kogishi K, Higuchi K, Matushita T, Wang J, Chiba T, Hosokawa M (2000) Serum lipid concentrations and mean life span are modulated by dietary polyunsaturated fatty acids in the senescence-accelerated mouse *J Nutr* 130:221–227
- Wang SJ, Kalpathy-Cramer J, Kim JS, Fuller CD, Thomas CR (2010) Parametric survival models for predicting the benefit of adjuvant chemoradiotherapy in gallbladder cancer *AMIA Annu SympProceed* 2010:847–851
- Williams GC (1957) Pleiotropy, natural selection and the evolution of senescence. *Evolution* 11:398–411
- Wilmoth JR, Horiuchi S (1999) Rectangularization revisited: variability of age at death within human populations. *Demography* 36:475–495
- Wilmoth JR, Deegan LJ, Lundström H, Horiuchi S (2000) Increase of maximum life-span in Sweden, 1861–1999. *Science* 289:2366–2368. <https://doi.org/10.1126/science.289.5488.2366>
- Wit J, Kristensen TN, Sarup P, Frydenberg J, Loeschcke V (2013) Laboratory selection for increased longevity in *Drosophila melanogaster* reduces field performance *Exper Gerontol* 48:1189–1195. <https://doi.org/10.1016/j.exger.2013.07.012>
- Witten M, Sätzer W (1992) Gompertz survival model parameters: Estimation and sensitivity *App Math Lett* 5:7–12. [https://doi.org/10.1016/0893-9659\(92\)90125-S](https://doi.org/10.1016/0893-9659(92)90125-S)
- Wu JJ, Liu J, Chen EB, Wang JJ, Cao L, Narayan N, Fergusson MM, Rovira II, Allen M, Springer DA, Lago CU, Zhang S, DuBois W, Ward T, deCabo R, Gavrilova O, Mock B, Finkel T (2013) Increased mammalian lifespan and a segmental and tissue-specific slowing of aging after genetic reduction of mTOR expression *Cell Reports* 4:913–920. <https://doi.org/10.1016/j.celrep.2013.07.030>
- Wu L, Zhou B, Oshiro-Rapley N, Li M, Paulo JA, Webster CM, Mou F, Kacergis MC, Talkowski ME, Carr CE, Gygi SP, Zheng B, Soukas AA (2017) An ancient, unified mechanism for metformin growth inhibition in *C. elegans* and cancer. *Cell* 167:1705–1718.e1713. <https://doi.org/10.1016/j.cell.2016.11.055>
- Yen K, Steinsaltz D, Mobbs CV (2008) Validated analysis of mortality rates demonstrates distinct genetic mechanisms that influence lifespan *Exper Gerontol* 43:1044–1051. <https://doi.org/10.1016/j.exger.2008.09.006>
- Zheng H, Yang Y, Land K (2012) Heterogeneity in the Strehler-Mildvan general theory of mortality and aging. *Demography* 48:267–290. <https://doi.org/10.1007/s13524-011-0013-8>
- Zwaan B, Bijlsma R, Hoekstra RF (1995) Direct selection on life span in *Drosophila melanogaster*. *Evolution* 49:649–659. <https://doi.org/10.2307/2410318>