ANNUAL REVIEWS

Annual Review of Statistics and Its Application

Is There a Cap on Longevity? A Statistical Review

Léo R. Belzile,[1] Anthony C. Davison,[2] Jutta Gampe,[3] Holger Rootzén,[4] and Dmitrii Zholud[4]

[1]Sciences de la décision, HEC Montréal, Montréal, Québec H3T 2A7, Canada; email: leo.belzile@hec.ca

[2]Institute of Mathematics, École polytechnique fédérale de Lausanne, 1015 Lausanne, Switzerland

[3]Laboratory of Statistical Demography, Max Planck Institute for Demographic Research, 18057 Rostock, Germany

[4]Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 41296 Gothenburg, Sweden

Annu. Rev. Stat. Appl. 2022. 9:21–45

First published as a Review in Advance on November 10, 2021

The *Annual Review of Statistics and Its Application* is online at statistics.annualreviews.org

https://doi.org/10.1146/annurev-statistics-040120-025426

Copyright © 2022 by Annual Reviews. All rights reserved

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

censoring, data validation, extreme old age, generalized Pareto distribution, Gompertz distribution, Lexis diagram, survival analysis, supercentenarian, truncation

Abstract

There is sustained and widespread interest in understanding the limit, if there is any, to the human life span. Apart from its intrinsic and biological interest, changes in survival in old age have implications for the sustainability of social security systems. A central question is whether the endpoint of the underlying lifetime distribution is finite. Recent analyses of data on the oldest human lifetimes have led to competing claims about survival and to some controversy, due in part to incorrect statistical analysis. This article discusses the particularities of such data, outlines correct ways of handling them, and presents suitable models and methods for their analysis. We provide a critical assessment of some earlier work and illustrate the ideas through reanalysis of semisupercentenarian lifetime data. Our analysis suggests that remaining life length after age 109 is exponentially distributed and that any upper limit lies well beyond the highest lifetime yet reliably recorded. Lower limits to 95% confidence intervals for the human life span are about 130 years, and point estimates typically indicate no upper limit at all.

1. INTRODUCTION

The possible existence of a hard upper limit, a cap, on human lifetimes is hotly debated and attracts widespread public interest. Answers to the question are important demographically and for pension systems, since the presence of a cap may imply that the upward trend in expected lifetimes that has occurred over the past century in developed countries cannot continue. Moreover, establishing the existence or lack of such a limit would inform biological theories of aging and could aid efforts to prolong lives.

In this article we discuss statistical approaches to understanding extreme human lifetimes and to investigating whether a limit exists. Several components to this interact: the availability and quality of data, appropriate statistical models and methods, and extrapolation. We illustrate these issues by reanalyzing data on extreme human lifetimes, give a critical review of earlier work, and present our view on the debate. Throughout, we use the term "(human) life span" for the generic, theoretical, limit to the lives of humans, and the term "lifetime" to refer to the life length of an individual human.

The first component, data sources and data quality, is discussed in Section 2. In statistical terms, our opening question—Is there a cap on longevity?—corresponds to investigating whether the support of the lifetime distribution is bounded or unbounded and, more practically, how this may be inferred from data. One key aspect is whether the available data are representative of the population under study. This is the first aspect discussed in Section 2. As the research question requires information on the right-hand tail of the lifetime distribution, reliable analyses depend critically on the quality of data on unusually long lifetimes. Extreme ages at death are well above 100 years, and although many developed countries have long, high-quality records of births and deaths, this implies that we must rely on data that go back more than a century. Continuous archiving may have been obstructed by wars or changes of territory or political systems, and even complete data series may be tarnished by errors of entry and transcription that compromise studies of very old individuals. Data validation is also addressed in Section 2.

In addition to representative sampling and data validation, a further complication is the effect of different observation schemes. Rounding, censoring, and truncation are very common with lifetime data; the first two reduce information on the observations, and the third determines which observations are available at all. In Section 3 we illustrate the effect of truncation, discuss methods for removing the resulting biases, and outline key lifetime data concepts. A central notion is the hazard of death, or force of mortality, as it is called in demography. This is the failure rate of the lifetime distribution; it quantifies the risk of death in the next infinitesimal interval, conditional on survival thus far. The hazard function for a continuous variable that has a finite upper bound must increase to infinity as the variable approaches that bound, a phenomenon sometimes very expressively called a wall of death. A hazard that increases to infinity, however, need not impose a finite life span, and this has led to some confusion. A lifetime distribution whose hazard function is bounded must have an infinite upper limit.

Section 3 also reviews the analysis of extreme lifetimes, including the use of the generalized Pareto distribution arising in extreme value statistics and the Gompertz distribution commonly used in demography, and discusses how heavy rounding, truncation, and censoring impact estimation of the hazard.

Most statistical analyses involve some form of extrapolation, from a sample to a population or from the past to the future. Statistical extrapolation entails giving some idea of the underlying uncertainty, formal statement of which requires a stochastic model. Semi- and nonparametric models are complex, flexible, and widely used in lifetime data analysis but are unreliable outside the range of the data, whereas parametric models are simpler and allow extrapolation,

though this may be biased if the model describes the data poorly. Here our focus is extrapolation beyond the data to a possible limiting life span, and this requires parametric modeling, but the fit of competing models must be carefully assessed, alongside the range of ages over which they are adequate.

There is no doubt that the force of mortality increases up to at least 105 years, so if one uses a parametric model to analyze datasets dominated by lifetimes about 100 years, the fitted model will reflect this increase. For the generalized Pareto model, this necessarily leads to a finite life span estimate, whereas for the Gompertz model, which is widely fitted at lower ages, the force of mortality increases indefinitely but the estimated life span is infinite. Such extrapolations were the only ones possible before the existence of the International Database on Longevity (IDL), but they agree badly with these data, which show a plateauing of the force of mortality. Hence, extrapolations from the age of about 100 to ages 120 or higher cannot now be regarded as trustworthy, and inferences far beyond the IDL data range of 105–122 years must be treated with circumspection.

In Section 4 we illustrate the consequences of extrapolation from ages around 100 and of rounding and truncation. Plateauing is a central issue at higher ages: Does the force of mortality become constant at some point, or does it continue to increase, perhaps so much that a finite life span is effectively imposed? We review the part of the literature that concludes there in fact is plateauing, perhaps around age 109, and note that this agrees with our analyses. We also review differences in mortality at extreme age between gender, between persons born earlier or later, and between countries, adding results for newly-available IDL data. The impression is that such differences are limited.

In Section 5 we discuss the analysis strategies of some of the articles that find evidence for a cap on the human life span. Key issues here are the handling of truncation and censoring and extrapolation from relatively low to much higher ages. Our overall conclusions are given in Section 6.

2. DATA ON EXTREME LIFETIMES

2.1. General

Data on extreme ages at death and the number living at those ages are used to investigate the upper tail of the lifetime distribution. Such data must be collected to reflect the actual tail of the lifetime distribution and not a distorted version of it. In statistics, this is termed representative sampling, and it requires a well-defined and correctly implemented observation scheme as well as a clear definition of the population being considered.

A key element is the age range that is deemed to contain the extreme lifetimes—that is, which ages form the upper tail. Supercentenarians, who survive to or beyond their 110th birthdays, are commonly considered to constitute this group. If we extend the age range downward by five years, then we may include semisupercentenarians, who live to see their 105th but not their 110th birthdays.

Information on potential (semi)supercentenarians should be treated with caution. Age overstatement is all too frequent, as a very long life is highly respected, so data on supercentenarians must be carefully and individually validated to ascertain that the reported age at death is correct. Such validation ideally requires a check of birth or baptism records, preferably supplemented by early-life documents, such as early census records or marriage certificates, that can be conclusively linked to the individual. The documents to be used depend on the national administrative system. Not surprisingly, age validation of foreign-born individuals is often difficult, so studies are frequently limited to native-born persons. Poulain (2010) discusses age-validation in detail.

2.2. Data Resources

There are two main resources for validated supercentenarian lifetimes: the IDL and the collection of the Gerontology Research Group. They can be found online at **https://www.supercentenarians.org** and at **https://grg.org/**.

The IDL contains data assembled by government agencies. The data collection varies by country but is always based on a well-defined and documented procedure that should be respected in any statistical analysis. This is to avoid age-ascertainment bias, which arises if the inclusion of an individual in the sample depends, in an unspecified way, on his or her age. Such bias is typically present if record-holders or individuals with extreme ages are more likely to appear in a dataset than younger ones. As this unequal sampling is not quantifiable, the resulting distortion cannot be managed. Typically the data are intended to be representative of all individuals who are born and die in the same country, and comprise all those found in the national records and successfully validated who die above a given age between two calendar dates. Section 3.2.1 explains a typical observation scheme in more detail.

The IDL was originally compiled to allow the estimation of the trajectory of human mortality, i.e., the hazard of death, up to the highest ages (Cournil et al. 2010). Initially the IDL provided data only on supercentenarians, but it now includes semisupercentenarians for some countries (Jdanov et al. 2021). At the time of writing, in spring 2021, it contains validated supercentenarian lifetimes from 13 countries and information on semisupercentenarians for nine countries. As it is updated often, care is needed when attempting to reproduce previous analyses.

The Gerontology Research Group website provides lists of known and validated supercentenarians, including some currently alive: on April 1, 2021, for example, the database stated that the oldest living validated person, Kane Tanaka, was aged 118 years and 94 days. It contains many interesting facts, but age-ascertainment bias and the absence of a specified observation scheme render it unsuitable for statistical analysis.

Certain other databases that qualify as ascertainment bias–free samples of long-lived individuals and have been used in lifetime analyses are discussed in Section 4, including one from the Italian National Statistics Institute (ISTAT) used in Sections 3.2 and 4.4.

3. STATISTICAL MODELING

Statistical analysis of extreme lifetimes rests on survival analysis and extreme value models. The first is discussed in Sections 3.1–3.3. We introduce basic concepts, discuss the censoring and truncation that certain observation schemes entail, and explain how they can be accommodated by correctly setting up the likelihood. Although we emphasize parametric models, we also briefly discuss nonparametric estimation (Section 3.1.3). We illustrate specific problems of observation schemes that have gone unnoticed in some previous research on life span limits in Section 3.2 and give an overview of the most popular models for (old-age) human mortality in Section 3.3. Section 3.4 presents threshold exceedance models from extreme value theory. The pivotal role of the generalized Pareto distribution is highlighted (Section 3.4.1), and penultimate approximations and their importance in the assessment of tail behavior are discussed in Section 3.4.2. Extended models are presented in Section 3.4.3.

3.1. Elements of Survival Analysis

Survival analysis began with John Graunt's (1662) actuarial life table and remains in full development. Systematic accounts may be found in books such as those by Cox & Oakes (1984), Andersen et al. (1993), Therneau & Grambsch (2000), Aalen et al. (2008) and Moore (2016).

3.1.1. Basic notions. Consider a positive continuous random variable T with distribution and density functions F and f, and let $t_F = \sup\{t : F(t) < 1\}$ denote the largest possible value of T, which may be finite or infinite. The survivor and hazard functions $S(t) = 1 - F(t)$ and $h(t) = f(t)/S(t)$ play central roles in survival analysis and satisfy

$$-\log S(t) = \int_0^t h(x)\,\mathrm{d}x = H(t),$$

where the cumulative hazard function H should tend to infinity as $t \to t_F$; if not, $S(t)$ does not tend to zero, implying that the event $T = t_F$ has a positive probability, sometimes called the cure fraction if $t_F = \infty$, in which case it corresponds to long-term survival. If t_F is finite and the cure fraction is zero, then $h(t)$ must increase sufficiently rapidly that $H(t) \to \infty$ as $t \to t_F$.

The hazard function $h(t)$, known in demography as the force of mortality, is the rate of failure in the infinitesimal interval $[t, t + \mathrm{d}t)$ conditional on survival to t. It plays a central role in survival analysis. The simplest possibility is that $h(t)$ equals a positive constant λ, and then T is exponentially distributed with mean λ^{-1}. Other possibilities are decreasing or increasing failure rate, whereby $h(t)$ is monotone decreasing or monotone increasing in t, or a convex bathtub shape.

Discrete survival data can arise when continuous responses are heavily rounded, and in nonparametric estimation of quantities defined for continuous data (see Section 3.1.3). In discrete time, failures are supposed to be possible at times $0 \le t_1 < t_2 < \cdots$ with positive probabilities $f_j = \Pr(T = t_j)$. In this case we define the hazard function by

$$h_j = \frac{f_j}{f_j + f_{j+1} + \cdots}, \quad \Pr(T > t_j \mid T \ge t_j) = 1 - h_j, \qquad 1.$$

can write

$$f_j = h_j \prod_{i=1}^{j-1}(1 - h_i), \quad S(t) = \prod_{i:t_i<t}(1 - h_i), \quad t_j < t \le t_{j+1}, \qquad 2.$$

and define the cumulative hazard function as $H(t) = -\sum_{i:t_i<t}\log(1 - h_i)$ for compatibility with the continuous case; when the h_i are small one may replace $-\log(1 - h_i)$ by h_i, giving the more natural definition $H(t) = \sum_{i:t_i<t} h_i$.

Maximum likelihood estimation and accompanying inferential tools such as likelihood ratio testing are standard in survival analysis. They provide a well-understood inferential framework that is readily adapted to allow for censoring and truncation, and can be regarded as approximations to Bayesian inferences with vague priors, which may be preferred by some investigators. We illustrate both likelihood and Bayesian methods below.

3.1.2. Truncation and censoring. Survival times are often subject to partial observation. The simplest form of this is right censoring, whereby T is supplemented with an indicator variable δ, with $\delta = 1$ if T is observed to take value t, and $\delta = 0$ if T is right-censored—i.e., it is known only that $T > t$. In this case the likelihood contribution can be written as $f(t)^\delta S(t)^{1-\delta} = h(t)^\delta S(t)$. Left censoring and interval censoring can also arise; a left-censored response is known only to satisfy $T < t$, whereas an interval-censored response is known only to fall into an interval $\mathcal{I}$ and would give a likelihood contribution $\Pr(T \in \mathcal{I})$. In the present setting, interval censoring could arise when the age at death in years, a, but not in days, is provided, so an individual has lifetime T falling into $\mathcal{I} = [a, a + 1)$ years.

Truncation determines which observations are available for analysis, whereas censoring reduces the information that can be extracted from them. Under truncation, a unit appears in the dataset only if its response variable falls into a subset, often an interval $\mathcal{I} = [a, b]$, of its possible values.

In this case the likelihood contribution is of the form $f(t \mid T \in \mathcal{I})$, with the conditioning expressing the restriction that T is available only if it lies in $\mathcal{I}$.

In the present context lifetimes are commonly interval-truncated, i.e., truncated to both left and right, or left-truncated and right-censored. The first arises when individuals are only included in a sample if they die within $\mathcal{I}$. The second arises when individuals who die within $\mathcal{I}$ or are alive at time b are included; only the age at time b is known for the latter group. The corresponding likelihood contributions may respectively be written as

$$\frac{f(t)}{S(a) - S(b)}, \quad a < t < b, \text{ and } \qquad \frac{h(t)^{\delta} S(t)}{S(a)}, \quad t > a, \tag{3.}$$

where the possible observations (t, δ) in the second case are $(b, 0)$ if t is right-censored at b and $(t, 1)$ if t is fully observed; then $a < t < b$ (see also Section 3.2).

Care is needed when analyzing the IDL lifetimes, since the truncation bounds for the semisupercentenarians and supercentenarians may differ (see Section 3.2.2).

3.1.3. Nonparametric estimation. Nonparametric estimates of the survivor and cumulative hazard functions cannot be used to extrapolate outside the observed data and so cannot directly address the issue of a cap on longevity, but they can be used to compare the fits of competing parametric models—for example, through quantile-quantile (Q–Q) plots that account for the observation scheme (Waller & Turnbull 1992). Nonparametric estimates can only place mass on the distinct observed failure times, as the data themselves contain no evidence that failure is possible at any other time.

Consider a random sample of survival times $t_1, \ldots, t_n$ from a distribution F, where it is supposed that $a_i < t_i < b_i$ and the truncation limits a_i and b_i equal zero and infinity if t_i is not truncated. With no censoring or truncation, the nonparametric maximum likelihood estimate $\widehat{F}$ of F places masses $1/n$ on each of the t_i, and the survivor and cumulative hazard functions can be estimated through Equations 1 and 2. Modification of $\widehat{F}$ to allow for left truncation and right censoring yields the Kaplan–Meier or product-limit estimator (Kaplan & Meier 1958, Tsai et al. 1987), which is most simply expressed in terms of the number of failures $d_j \in \{0, \ldots, r_j\}$ at the distinct failure times $t'_1 < \cdots < t'_J$ and the number r_j of individuals still at risk, i.e., all individuals not yet failed or censored at t'_j. Then $\widehat{h}_j = d_j / r_j$ and

$$\widehat{S}(t) = \prod_{j:t'_j<t} (1 - \widehat{h}_j), \quad \operatorname{var}\left\{\widehat{S}(t)\right\} \doteq \widehat{S}(t)^2 \sum_{j:t'_j<t} \frac{\widehat{h}_j}{r_j(1 - \widehat{h}_j)}, \quad t > 0.$$

Andersen et al. (1993, section IV) review methods for obtaining pointwise and simultaneous confidence intervals for nonparametric estimators of the hazard and survivor functions.

There are no general explicit formulae for estimators under more complex observation schemes. One approach for right-censored and interval-truncated data applies the expectation–maximization (EM) algorithm (Dempster et al. 1977). Let the random variable T_i taking value t_i have truncation set $\mathcal{T}_i \subset \{1, \ldots, J\}$; this is the set of j for which $T_i = t'_j$ is possible. Likewise, define a censoring set $\mathcal{C}_i \subset \mathcal{T}_i$ containing all the indices j for which $T_i = t'_j$ is possible; if T_i is observed then $\mathcal{C}_i$ contains a single element, but if T_i is censored then $\mathcal{C}_i$ has the form $\{j_1, \ldots, j_2\}$. We associate indicator variables γ_{ij} and τ_{ij} to the censoring and truncation sets: $\gamma_{ij} = 1$ if and only if $j \in \mathcal{C}_i$ and $\tau_{ij} = 1$ if and only if $j \in \mathcal{T}_i$, for $i = 1, \ldots, n$ and $j = 1, \ldots, J$. If the T_i can only take values in $t'_1, \ldots, t'_J$, then nonparametric maximum likelihood estimation of F involves finding the

probabilities $f_j = \Pr(T = t'_j) \geq 0$ that maximize the likelihood

$$\prod_{i=1}^{n} \frac{\sum_{j=1}^{J} \gamma_{ij} f_j}{\sum_{j=1}^{J} \tau_{ij} f_j} \qquad 4.$$

subject to $\sum_{j=1}^{J} f_j = 1$ (Turnbull 1976). The likelihood might be maximized directly if J is very small, but it is typically preferable to note that if the T_i had been observed without censoring or truncation then the log-likelihood written in terms of indicator variables $I_{ij} = I(t_i = t'_j)$ would be $\sum_{i=1}^{n} \sum_{j=1}^{J} I_{ij} \log f_j$, yielding $\widehat{f}_j = \sum_i I_{ij} / \sum_{i,j} I_{ij}$; this is the maximization (M) step of an EM algorithm. The expectation (E) step involves replacing I_{ij} at the mth iteration by

$$\widehat{I}_{ij}^{m} = \frac{\gamma_{ij} \widehat{f}_j^{m}}{\sum_{j'=1}^{J} \gamma_{ij'} \widehat{f}_{j'}^{m}} + \frac{(1 - \tau_{ij}) \widehat{f}_j^{m}}{\sum_{j'=1}^{J} \tau_{ij'} \widehat{f}_{j'}^{m}}, \quad i = 1, \ldots, n, \; j = 1, \ldots, J, \qquad 5.$$

where $\widehat{f}_1^{m}, \ldots, \widehat{f}_J^{m}$ result from the M step. The two terms in Equation 5 correspond to the probability that a possibly censored T_i equals t'_j and to ghosts, i.e., individuals who would have been observed had there been no truncation. The algorithm consists of setting $\widehat{f}_1^{0} = \cdots = \widehat{f}_J^{0} = 1/J$, computing the $\widehat{I}_{ij}^{1}$ by the E step and then $\widehat{f}_1^{1}, \ldots, \widehat{f}_J^{1}$ by the M step, increasing m, and then alternating the E and M steps until convergence. The observed information is found by differentiating the logarithm of Equation 4.

3.1.4. Diagnostic plots. Standard graphical diagnostics must be modified to account for different observation schemes. If F_0 is an estimated or postulated parametric distribution and F_n is a nonparametric estimator of the distribution function, we can construct a Q–Q plot by plotting observed failure times t_i against the positions $v_i = F_0^{-1}\{F_n(t_i)\}$ on the x-axis. With truncated data, $a_i < t_i < b_i$, so each observation has a different distribution, yielding $v_i = F_0^{-1}[F_0(a_i) + \{F_0(b_i) - F_0(a_i)\}\{F_n(t_i) - F_n(a_i)\}/\{F_n(b_i) - F_n(a_i)\}]$. One effect of truncation is that the ranks of v_i and t_i need not coincide, so the graph need not be monotone. Approximate confidence intervals may be obtained by parametric bootstrapping (Belzile et al. 2021) (see the **Supplemental Appendix**).

Supplemental Material >

3.2. Observation Schemes and Likelihoods

Appropriate specification of the likelihood requires a correct representation of how the available data came to be observed, since subsequent inferences will otherwise be unreliable. We now describe some standard observation schemes for data on extreme lifetimes and the construction of the corresponding likelihoods.

3.2.1. Basic notions. Creating a sample for analysis of extreme ages usually starts by casting a net (Kestenbaum & Ferguson 2010) on the population of potential (semi)supercentenarians. The candidates identified in this step are then individually validated. The population on which the net is cast comprises, for the most part, deceased individuals who died above a certain age u_0, here 110 or 105 years. The specific observation scheme depends on how the net is defined.

One popular design is the following: All individuals who died at or above age u_0 between two calendar dates c_1 and c_2 are collected and validated. Individual excess lifetimes above u_0 are depicted in a Lexis diagram (Keiding 1990), in which each individual lifetime is shown as a diagonal line of unit slope, commencing at a calendar time x on the horizontal axis and later ending in death or censoring at calendar time $x + t$. In this article we are concerned only with lifetimes exceeding a threshold u_0, so only the part of the diagram containing the excess lifetimes above u_0 is shown. **Figure 1*a*** illustrates the effect of this observation scheme for lifetimes sampled from an exponential distribution; only individuals with dark blue lines are retained.

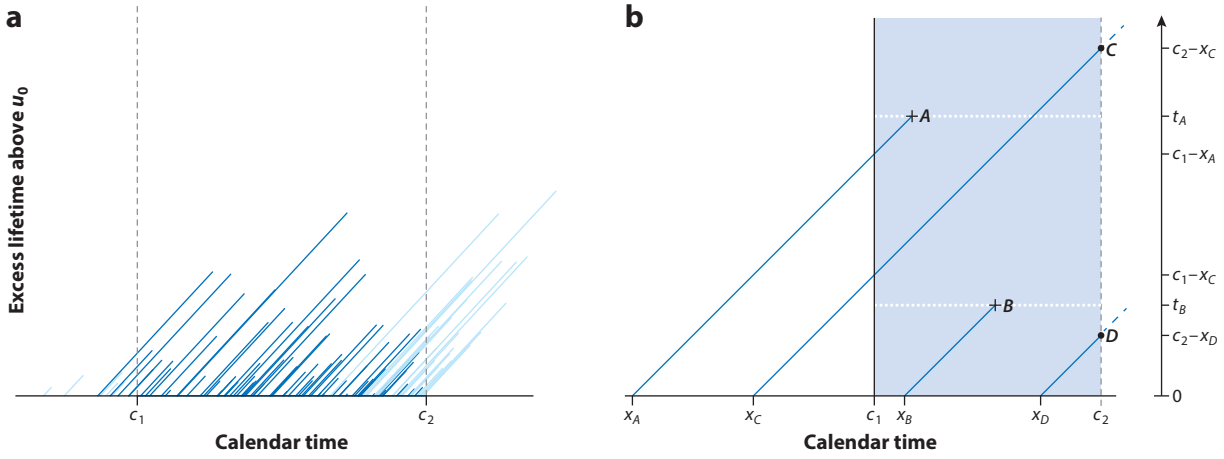

Figure 1

Lexis diagrams and observation schemes. (*a*) Exponential life trajectories sampled under left and right truncation; only individuals with dark blue trajectories appear in the sample. (*b*) Schematic showing life trajectories of individuals subject to left truncation and right censoring as in the ISTAT data; those who would attain age u_0 after calendar time c_2 or who died before calendar time c_1 or below age u_0 are unobserved. Individual A passes age u_0 at date $x_A < c_1$, has excess lifetime of $c_1 - x_A$ at calendar time c_1, and dies at age $u_0 + t_A$; she is left-truncated at excess lifetime $c_1 - x_A$. Individual B reaches age u_0 at calendar time x_B and dies between calendar times c_1 and c_2 at age $u_0 + t_B$. Individual C is left-truncated at excess lifetime $c_1 - x_C$ and right-censored at excess lifetime $c_2 - x_C$. Individual D is right-censored at excess lifetime $c_2 - x_D$. If the data were right-truncated, then individuals C and D, who do not die in the shaded region $\mathcal{C} = [c_1, c_2] \times [0, \infty)$, would not appear in the sample.

Conclusions may be biased if the analysis does not take the observation scheme into account. For the truncated data depicted in **Figure 1*a***, for death occurring close to c_1, only individuals who survived beyond the age they attained at time c_1 are included, so longer lifetimes are overrepresented. Furthermore, only individuals who died before time c_2 appear in the sample, leading to an overrepresentation of shorter lifetimes from near c_2. In many countries the number of individuals living beyond age u_0 is growing over time. The data were simulated to reflect this, and the resulting preponderance of shorter lifetimes from near c_2 over the longer ones from near c_1 biases the sample downwards; **Figure 2*a*** compares the observed excess lifetimes with the true lifetime distribution. Truncation also leads to a downward trend in the maximum age achieved in the later birth cohorts, as in **Figure 2*b***. Such distortions can be removed if the observation scheme is known (see Sections 3.1.2 and 3.2.3).

3.2.2. International Database on Longevity observation frame. The IDL data available from **https://www.supercentenarians.org/** in March 2021 only include records of dead individuals and thus do not match exactly the description in Jdanov et al. (2021); data for Switzerland and Italy have also been removed for confidentiality reasons.

Recent versions of the IDL include both semisupercentenarians who died between calendar dates d_1 and d_2 and supercentenarians who died between calendar dates c_1 and c_2. The sampling frame thus consists of two regions, $[d_1, d_2] \times [105, 110)$ and $[c_1, c_2] \times [110, \infty)$, with potentially different truncation bounds. This complicates the analysis because many lifetimes are doubly interval-truncated, as described in more detail in the next paragraph, and handling the constraints requires a careful case-by-case analysis. Inclusion criteria can be deduced from the Lexis diagram by considering the birth dates of all individuals of a given country. All countries in the IDL database

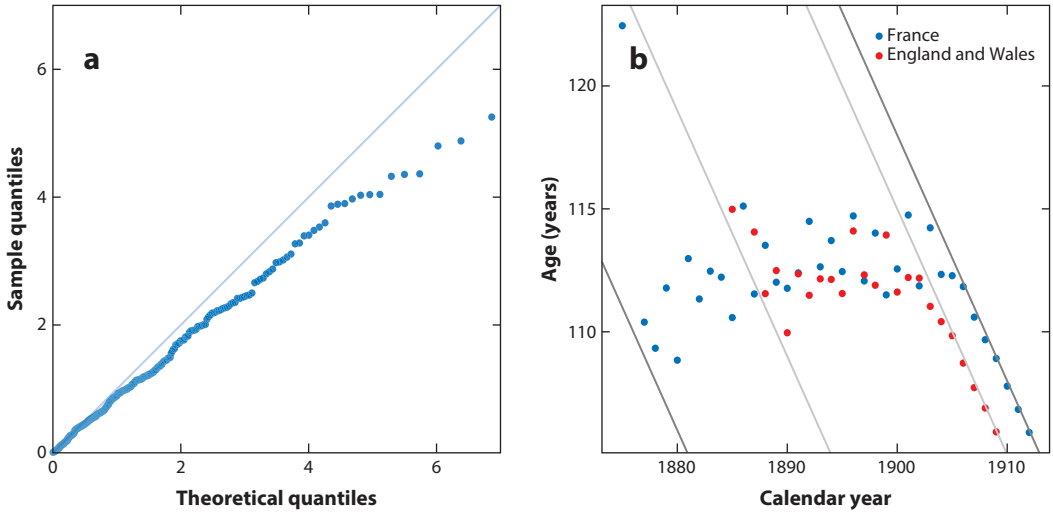

Figure 2

Effects of left and right truncation. (*a*) Quantile-quantile plot of the truncated exponential lifetimes y_i shown in **Figure 1*a*** against exponential plotting positions. (*b*) Maximum reported age at death of International Database on Longevity individuals for France and for England and Wales by birth cohort. The sloping lines show the minimum and maximum possible reported ages given the sampling intervals. As the individuals must be aged at least 105, the maximum age at death for the French birth cohort 1910, say, is just under 108, since for these data c_2 corresponds to 31 December 2017.

with semisupercentenarians have $c_2 > d_1$, and we assume this hereafter. Let x^{110} and x^{105} denote the calendar dates at which an individual reaches 110 and 105 years; **Figure 3** shows a typical IDL sampling frame with some scenarios discussed next.

If we were only considering supercentenarians, lifetimes would typically be left-truncated by the age reached at $\max(x^{110}, c_1)$ and right-truncated at c_2. However, two instances can lead to double interval truncation. The first can occur when $d_1 < c_1$; if $x^{105} < d_1 < x^{110} < c_1$, as for trajectory A, the trajectory is observed in $[d_1, x^{110}] \cup [c_1, c_2]$. The second can occur when $d_2 < x^{110} < c_2$, as for trajectory F, in which case some lifetimes lie in $[\max(d_1, x^{105}), d_2] \cup [x_{110}, c_2]$. If $d_1 < x^{110} < d_2$, some individuals could have traversed the set $[d_1, d_2] \times [105, 110)$ even when they died beyond 110 years, so their minimum observable age is that reached at $\max(x^{105}, d_1)$. Similarly, people who died in [105, 110) could have become supercentenarians and seen their death recorded if their 110th birthday x^{110} lies in $[c_1, c_2]$; then their lifetime is right-truncated by the age they would have reached at c_2. For all other individuals in [105, 110), the maximum observable age is capped at that reached at $\min(d_2, x^{110})$. If $x^{110} > c_2$ and we observe death in $[d_1, d_2] \times [105, 110)$, then the maximum possible age achieved is $\min(d_2, x^{110})$. **Figure 3*a*** illustrates the case when $d_1 < c_1$, whereas **Figure 3*b*** shows an example where $d_2 < c_2$.

This discussion underscores the importance of accurate metadata, the absence of which precludes correct inference.

3.2.3. Likelihood contributions. We now discuss how likelihood contributions depend on the underlying observation scheme. Consider **Figure 1*b***, in which the trajectory of each individual lifetime is shown as a diagonal line of unit slope. Let $\mathcal{C}$ denote the region $[c_1, c_2] \times [u_0, \infty)$, where the calendar times c_1 and c_2 are the sampling limits. Individuals whose trajectories do not intersect $\mathcal{C}$ are unobserved; indeed, their existence cannot be inferred from the available data.

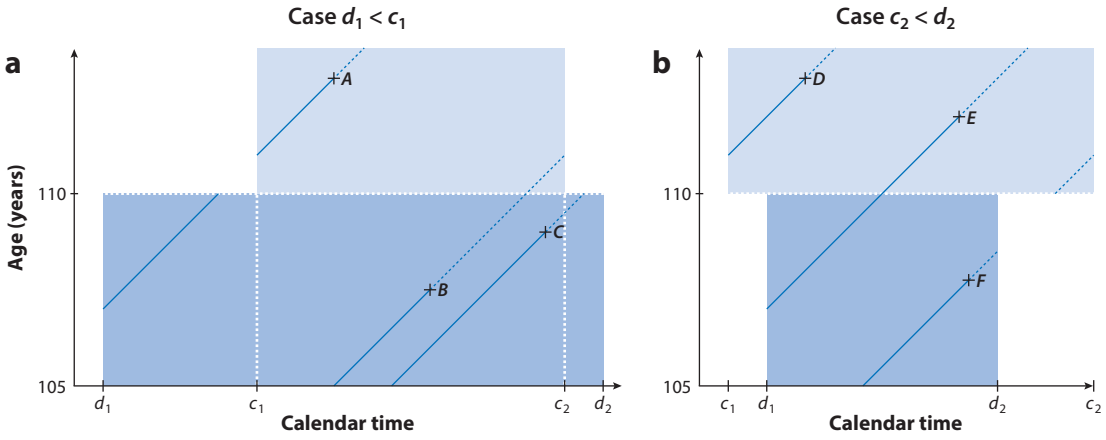

Figure 3

Lexis diagrams of the IDL data for countries with semisupercentenarian records, consisting of a sampling frame defined as regions in $[d_1, d_2] \times [105, 110)$ and $[c_1, c_2] \times [110, \infty)$. The full lines give the life trajectories until death (+), with dotted lines indicating the observation window. (*a*) Case $d_1 < c_1$: The lifetime trajectory of individual A is observed above age 105 in $[d_1, x_A^{110}] \cup [c_1, c_2]$, that of individual B in $[x_B^{105}, c_2]$, and that of individual C in $[x_C^{105}, x_C^{110}]$. (*b*) Case $c_2 > d_2$: The lifetime trajectory of individual D is observed in the interval $[c_1, c_2]$, that of individual E in $[d_1, c_2]$, and that of F in $[x_F^{105}, d_2] \cup [x_F^{110}, c_2]$.

Let $\mathcal{T} = \{(x + s, u_0 + s) : s > 0\}$ be the trajectory of someone who passes age u_0 at calendar time x, and let T denote her excess lifetime after u_0. For example, if u_0 corresponds to 105 years, then x would be the calendar date of her 105th birthday, and if her excess lifetime took value $T = t$, then she would trace the line from (x, u_0) to $(x + t, u_0 + t)$.

If there is interval truncation, then the argument in Section 3.1.2 implies that the appropriate likelihood contribution for a person with trajectory $\mathcal{T}$ who dies inside $\mathcal{C}$ is $f(t \mid T \in \mathcal{I})$, where $\mathcal{I}$ denotes the projection of $\mathcal{C} \cap \mathcal{T}$ onto the vertical axis of a Lexis diagram (**Figure 1*b***), and it is then easy to check that one should take

$$\mathcal{I} = [a, b] = [\max(0, c_1 - x), c_2 - x] \qquad 6.$$

in the first expression in Equation 3. If there is left truncation and right censoring, as in the ISTAT data analyzed in Section 4.4, then these values of a and b are inserted into the second expression of Equation 3. Hence, each individual independently contributes a conditional density term to the overall likelihood.

Truncation may be hidden. Suppose, for example, that annual cohorts defined by some suitable criteria and labeled by the years $y_{\min}, \ldots, y_{\max}$ are available, but only those for years $y_{\min}, \ldots, y^*$ are retained for analysis because some members of cohort $y^* + 1$ are still alive at time c_2. As only extinct cohorts are used, truncation and censoring appear to be absent. This would simplify analysis if it were true, but truncation is present nonetheless. The boundary c_1 in the Lexis diagram is effectively $-\infty$, but y^* has been chosen using the survival times, as the corresponding random variable Y represents the last cohort in which no one dies after c_2. Suppose the persons in cohorts $y_{\min}, \ldots, y^*$ enter $\mathcal{C}$ at dates $x_1, \ldots, x_n$ and have lifetimes $T_1, \ldots, T_n$ above u_0. Then,

$$\Pr(Y = y^*) = \prod_{j=1}^{n} \Pr(T_j \le c_2 - x_j) \times p,$$

where p is the probability that at least one individual in cohort $y^* + 1$ is alive at c_2. The joint probability element for the observed lifetimes $t_1, \ldots, t_n$ is $\prod_{j=1}^{n} f(t_j) \times p$, and conditioning on the

selection event $Y = y^*$ yields

$$f(t_1, \ldots, t_n \mid Y = y^*) = \prod_{j=1}^{n} \frac{f(t_j)}{F(c_2 - x_j)}, \quad 0 < t_j < c_2 - x_j, \quad j = 1, \ldots, n,$$

which is a product of terms for truncated observations in Equation 3, with $a_j = 0$ and $b_j = c_2 - x_j$. Thus, this approach entails truncation, though ignoring it may be harmless if $F(c_2 - x_j) \approx 1$ for all j. As using only extinct cohorts does not greatly simplify analysis, and the resulting smaller sample leads both to more variable estimates and to reduced power for model comparison and assessment, it cannot be recommended. The **Supplemental Appendix** illustrates the bias that can result from ignoring the truncation.

Supplemental Material >

It is superficially appealing to analyze the excess lifetimes for those persons dying in each year, ignoring the truncation, but doing so conflates the lifetime distribution $f(t)$ and the rate $\nu(x)$ at which individuals reach age u_0. The excess lifetime for an individual dying in the interval $[c_1, c_2]$ has density $f_C(t)$ proportional to $f(t)w(t)$, where the weighting function

$$w(t) = \int_{c_1 - t}^{c_2 - t} \nu(x)\, dx, \quad t > 0,$$

is decreasing because medical and social advances have increased ν over a long period. Hence, naive analysis of the excess lifetimes of the persons dying in C will yield an estimate of $f_C(t)$, which is a tilt of $f(t)$ toward the origin (see **Figure 2*a***).

In addition to taking the observation scheme into account, the use of conditional likelihood contributions such as those in Equation 3 has the important advantage that ν does not influence the analysis. Keiding (1990) and Davison (2018) provide more discussion.

3.3. Models for Human Lifetimes

The trajectory of age-specific mortality, i.e., the hazard of human lifetime, has proved to be remarkably stable in its overall shape but has dropped markedly over time (Burger et al. 2012). Parametric modeling of human mortality over the full age range is difficult due to its complex shape, but when the focus is on so-called senescent mortality, starting from mid-life, an exponential increase in human death rates was noticed by Gompertz (1825), leading to the hazard

$$h(t) = \sigma^{-1} e^{\beta t/\sigma}, \quad t > 0, \quad \beta, \sigma > 0, \tag{7.}$$

with survivor function

$$S(t) = \exp\left\{-(e^{\beta t/\sigma} - 1)/\beta\right\}.$$

This model is typically used to describe the increase of death risks with age for ages t over 50 years. Makeham (1860) added an age-independent component $\lambda > 0$ to give what is now called the Gompertz–Makeham model,

$$h(t) = \lambda + \sigma^{-1} e^{\beta t/\sigma}, \qquad H(t) = \lambda t + (e^{\beta t/\sigma} - 1)/\beta, \tag{8.}$$

which often proves better at lower ages. The dimensionless parameter β determines how fast the age-specific component increases; if $\beta = 0$, then this risk is constant and only the sum of the hazards, $\lambda + 1/\sigma$, is identifiable. Although its hazard function increases indefinitely for $\beta > 0$, the Gompertz–Makeham model imposes no finite upper limit on lifetimes. Taking $\beta < 0$ in Equation 7 yields a defective distribution, as then $S(t) \to \exp(1/\beta)$ as $t \to \infty$.

Perks (1932, p. 13) noted that "the Makeham curves ran much too high at the older ages," a phenomenon now known as late-life mortality deceleration: The exponential increase in death

rates slows down at high ages, a feature typically noticeable after about age 80 (Thatcher et al. 1998). Several extensions of the Gompertz–Makeham model have been proposed to capture this phenomenon (Thatcher 1999, Bebbington et al. 2011), the most prominent being a model where the exponential Gompertz component in Equation 8 is replaced by a logistic function. In its most general form, the hazard function is

$$h(t) = \lambda + \frac{Ae^{\beta t/\sigma}}{1 + Be^{\beta t/\sigma}}, \qquad A > 0,\ B \geq 0. \tag{9.}$$

Several special models can be obtained from Equation 9, including the Gompertz, Gompertz–Makeham, and gamma-Gompertz–Makeham models; in the last, individuals share a Gompertz component with the same age increase, i.e., the same value of β/σ, but hazard levels differ between individuals. If these are modeled by a gamma-distributed multiplicative random effect, the resulting marginal hazard is a logistic function (Beard 1971, Vaupel et al. 1979). A logistic hazard (Equation 9) increases exponentially for small t but ultimately reaches a plateau as $t \to \infty$.

3.4. Extreme Value Models

The presence or not of a cap on longevity can be viewed through the lens of extreme value statistics. The stochastic behavior of maxima and related quantities is well understood and is described from different viewpoints by Embrechts et al. (1997), Coles (2001), Beirlant et al. (2004), de Haan & Ferreira (2006), and Resnick (2006).

3.4.1. Threshold exceedances. A flexible approach to modeling high values of a random variable X with continuous distribution function F is to consider its exceedances $X - u$ above a threshold u (Pickands 1975). If, as u increases to the upper support point t_F of X, there exists a positive function a_u of u such that the rescaled exceedances converge to a nondegenerate distribution, then

$$\Pr\left\{(X-u)/a_u > t \mid X > u\right\} \to 1 - G(t) = \begin{cases} (1+\xi t/\sigma)_+^{-1/\xi}, & \xi \neq 0, \\ \exp(-t/\sigma), & \xi = 0, \end{cases} \tag{10.}$$

where $c_+ = \max(c, 0)$ for a real number c. Hence, the generalized Pareto distribution $G(t)$ provides a suitable statistical model for exceedances $T = X - u$ over a high threshold u. This distribution has a scale parameter σ, which depends on u, and a shape parameter ξ; T takes values in $(0, \infty)$ if $\xi \geq 0$ and in $(0, -\sigma/\xi)$ if $\xi < 0$. Its hazard function, $(\sigma + \xi t)_+^{-1}$, is constant for $\xi = 0$, declines slowly if $\xi > 0$, and increases without limit as $t \to t_F$ if $\xi < 0$. With threshold u and a negative shape parameter, the upper limit for human lifetimes would therefore be $\psi = u - \sigma/\xi$ under this model. A generalized Pareto variable with scale parameter σ is threshold-stable: For $v > 0$ and provided $\sigma_v = \sigma + \xi v$ is positive, a positive exceedance $T - v$ is also generalized Pareto, with scale parameter σ_v and shape parameter ξ. This result is useful in choosing the threshold u, since it gives stability relations that should be satisfied if the generalized Pareto model is adequate above u. Statistical aspects of this model are discussed by Davison & Smith (1990).

The limiting generalized Pareto survivor function, Equation 10, should provide a good approximation for data above a sufficiently high threshold u, so a key element in applications is the choice of u. Taking u too low risks the introduction of bias because the model provides a poor approximation to the distribution of exceedances, whereas taking u too high may winnow the sample so much that subsequent inference becomes too uncertain to be informative. Many approaches to choosing u have been proposed, including formal methods such as tests of fit and informal methods such as graphical assessment of whether the stability relations are satisfied (Scarrott & MacDonald 2012, Wadsworth 2016, Bader et al. 2018).

In Section 4 we consider Bayesian analyses that entail specification of prior distributions. The use of objective priors is particularly important when studying the endpoint of the generalized Pareto distribution, and we apply the maximal data information prior (Zellner 1977), which maximizes the information from the data relative to that from the prior. For the exponential and generalized Pareto distributions, this reduces to the priors σ^{-1} and $\pi(\sigma, \xi) \propto \sigma^{-1}\exp(-\xi - 1)$, respectively. The posterior density for the latter is improper unless the prior is truncated; we follow Northrop & Attalides (2016) and restrict the range to $\xi \geq -1$. Moala & Dey (2018) discuss the validity of the prior for the Gompertz distribution, $\pi(\sigma, \beta) \propto \sigma^{-1} \exp\{\exp(1/\beta) \int_{1/\beta}^{\infty} \exp(-t)/t\,\mathrm{d}t\}$, for a different parametrization. In these low-dimensional settings we can use the ratio-of-uniforms method (Wakefield et al. 1991, Northrop 2021) to sample from the posterior distribution.

3.4.2. Penultimate approximation. The shape parameter determines the behavior of the upper tail of the distribution of threshold exceedances, and under mild conditions this is determined by the behavior of the reciprocal hazard function $r(t) = 1/h(t)$. If this function has a continuous derivative r' and if we write $\xi_u = r'(u)$, then $\xi = \lim_{u \to t_F} \xi_u$, but it can be shown that for $u < t_F$ the generalized Pareto distribution with parameter ξ_u provides a better, so-called penultimate, approximation to the threshold exceedances (Smith 1987). The Gompertz–Makeham model, for example, has $t_F = \infty$ and

$$\xi_u = -\frac{\beta e^{-\beta u/\sigma}}{(1 + \lambda\sigma e^{-\beta u/\sigma})^2}, \quad u > 0,$$

so $\xi_u \to \xi = 0$ fairly rapidly as $u \to \infty$. Thus, we would expect threshold exceedances from Equation 8 for a sufficiently high u to be well approximated by an exponential distribution, stemming from Equation 10 with $\xi = 0$, but for somewhat lower thresholds a better approximation would be given by $\xi_u < 0$, yielding a model with a finite endpoint ψ. Hence, if the Gompertz–Makeham model was appropriate at all ages but a generalized Pareto distribution was fitted to exceedances of u, then as u increases we would expect to find that the estimates of ξ are initially negative but then approach zero.

One way to assess the penultimate behavior of F is to fit a model proposed by Northrop & Coleman (2014) in which generalized Pareto models with shape parameters $\xi_1, \ldots, \xi_K$ are fitted to data falling between thresholds $u_1 < \cdots < u_K < u_{K+1} = \infty$. In order for the resulting density to be smooth, the scale parameters $\sigma_1, \ldots, \sigma_K$ over the successive intervals should satisfy $\sigma_k = \sigma_1 + \sum_{k'=1}^{k-1} \xi_{k'}(u_{k'+1} - u_{k'})$ for $k = 2, \ldots, K$. Thus, the model has parameters $\sigma_1, \xi_1, \ldots, \xi_K$, which are estimated by treating the observations as independent, with those in the interval $(u_k, u_{k+1}]$ modeled using the truncated generalized Pareto distribution with shape ξ_k and scale σ_k.

3.4.3. Extended models. The generalized Pareto approximation for threshold exceedances holds under rather mild conditions on the underlying distribution F, but it may be useful to extend it by adding further parameters, in the hope of obtaining better fits in finite samples (Papastathopoulos & Tawn 2013).

Certain models simplify when certain parameters take boundary values. For example, comparison between the Gompertz and exponential distributions based on Equation 7 amounts to setting $\beta = 0$, which lies on the boundary of the parameter space, so standard large-sample results for likelihood ratio tests do not apply; the asymptotic significance level for a positive likelihood ratio statistic w_{Gom} is $\frac{1}{2}\Pr(\chi_1^2 > w_{\mathrm{Gom}})$, one-half of its usual value, and the significance level for $w_{\mathrm{Gom}} = 0$ is unity. Limiting distributions such as these can give poor finite-sample approximations, so simulation from the fitted null model is generally preferable for model comparison. **Table 1** illustrates this for the England and Wales semisupercentenarian data discussed in Section 4.3. The

Table 1 Likelihood ratio tests comparing Gompertz and exponential models for the England and Wales data

u	n_u	w_{Gom}	p_{b}	p_{a}	w_{GPD}	p_{GPD}
108	510	7.98	0.004	0.002	7.43	0.006
109	225	0.98	0.198	0.161	1.08	0.298
110	85	2.70	0.074	0.050	2.71	0.099
111	39	0.81	0.274	0.184	0.81	0.368

Shown are the range of thresholds u, with the number of exceedances n_u, the likelihood ratio statistics w_{Gom}, the bootstrap p-values p_{b}, and the asymptotic $\frac{1}{2}\chi^2_1$ p-values p_{a} for the null hypothesis $\beta = 0$. Also shown are the corresponding likelihood ratio statistics w_{GPD} for comparison of the generalized Pareto and exponential models, with their asymptotic significance levels p_{GPD}.

p-values obtained by simulating 10,000 datasets from the fitted exponential model with the same truncation bounds are systematically larger than the asymptotic p-values, because the asymptotic probability that $w_{\text{Gom}} = 0$ can be far from its finite-sample value.

4. DATA ANALYSES

We now reanalyze various existing datasets, both to illustrate how the various problems raised previously can undermine conclusions and to summarize results used in Section 6 to shed light on the question posed by the title of this article.

4.1. The Netherlands

We first demonstrate the presence of penultimate effects that could result in underestimation of the life span due to unreliable extrapolation. The data, from Einmahl et al. (2019), concern individuals who died aged at least 92 years between 1986 and 2015 and are from a Dutch population register; they are unvalidated and may be affected by age-ascertainment bias. There are lifetimes in days and the months and years of birth and death for 304,917 individuals, plus 226 persons whose lifetimes are unknown and so are interval-censored. All of these lifetimes are interval-truncated, and we treat them as such in our analysis. There is no sign of gender or cohort effects from age 105 onwards.

If the generalized Pareto distribution were a good approximation to lifetimes above age u, then estimates of its shape parameter ξ would be roughly constant for fits using higher thresholds. **Figure 4*a***, however, shows that these estimates increase steadily as u increases, suggesting that the lowest reasonable threshold here is 102 years.

For a deeper analysis we fit the Northrop & Coleman (2014) piecewise generalized Pareto model, handling truncation and censoring as described in Equation 3. With $K = 3$ and thresholds $u_1 = 98$, $u_2 = 101$, $u_3 = 104$, and $u_4 = 107$ years, the estimated shape parameters show a steady increase that is compatible with penultimate effects. This partially explains the results of Einmahl et al. (2019), whose fits are dominated by the lower ages.

The piecewise model can be used to check the hypothesis of equal shape through likelihood ratio tests: If $\xi_k = \cdots = \xi_K$, then the model is generalized Pareto above u_k with parameters (σ_k, ξ_k). Comparison of the piecewise model with thresholds $u_1, \ldots, u_4$ and a generalized Pareto model above u_k leads to rejection of the latter at thresholds 98 and 101. For threshold 104, the p-value is 0.49, so the two models appear to fit equally well.

We can also fit the Gompertz model to exceedances over a range of thresholds. The maximized log-likelihoods for the Gompertz and generalized Pareto models are equivalent from threshold

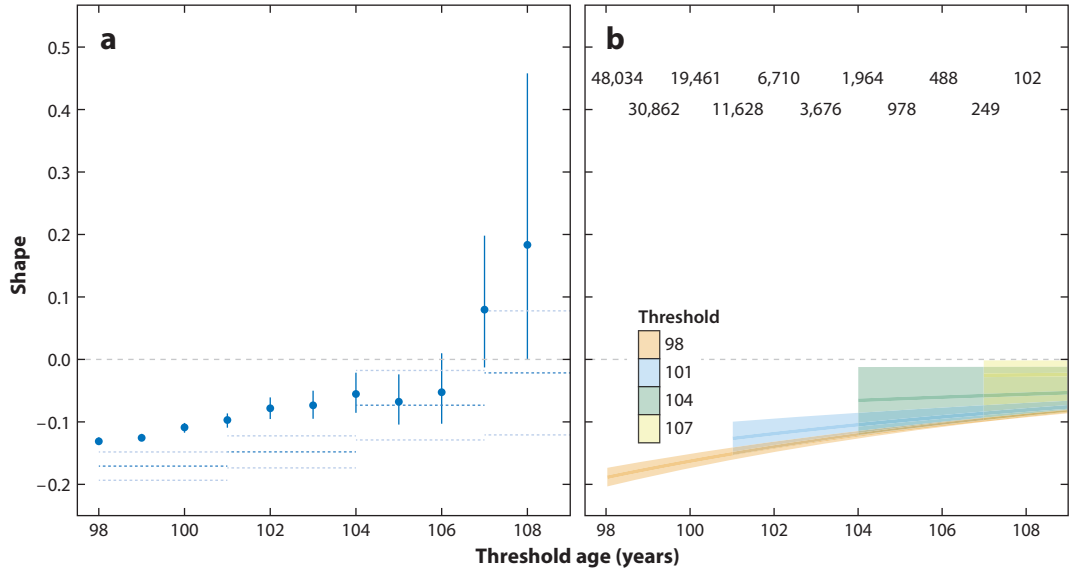

Figure 4

Penultimate effects in the Dutch data. (*a*) Maximum likelihood estimates of the shape parameter with thresholds 98, . . . , 108 years (*dots*) and 95% profile likelihood ratio confidence intervals and Northrop–Coleman estimates (*dashed dark blue lines*) with thresholds 98, 101, 104, and 107 and 95% confidence intervals (*dashed light blue lines*). (*b*) Estimated penultimate shape parameter ξ_u (*thick lines*) for the Gompertz model fitted to exceedances over thresholds 98, 101, 104, and 107 years, with 95% pointwise confidence regions (*shaded areas*). The numbers of exceedances above the yearly thresholds are reported in panel *b*.

102 onwards, and the models seem very similar until threshold 106. However, the Gompertz model cannot capture a decreasing hazard function; it reduces to an exponential model above high thresholds, which seems implausible based on the generalized Pareto fits at thresholds 107 and 108. If the Gompertz model were adequate, we would expect its tail behavior, given by the penultimate shape ξ_u obtained from fitted models, to broadly agree with the generalized Pareto shape estimates. **Figure 4*b*** shows 95% pointwise intervals for the curves corresponding to the penultimate shape ξ_u of the fitted Gompertz models at thresholds 98, 101, 104, and 107. These curves agree with the generalized Pareto shape estimates in the range 104–107 years, but not elsewhere; in particular, the shape parameter estimates at 107 seem to be positive, but those for the Gompertz model are capped at zero. The Gompertz model does not seem to be flexible enough to capture the penultimate properties of the data.

Validation of these data might change the results, particularly at the highest thresholds, but it appears that penultimate effects can be detected in these data until age 104 years. Although the reduced sample sizes render them undetectable, it seems plausible that such effects persist above age 104. Extrapolation is therefore unreliable if based on a generalized Pareto fit to individuals aged below 105 years and, if possible, a higher threshold is to be preferred. The inflexibility of the Gompertz hazard function suggests that it should be avoided for fits at very high ages.

4.2. Japan

Although we have stressed the central role of interval truncation, other types of data coarsening arise. We illustrate the effect of interval censoring and the asymmetric nature of inferences on the life span using data from Hanayama & Sibuya (2016), who analyze annual death counts since

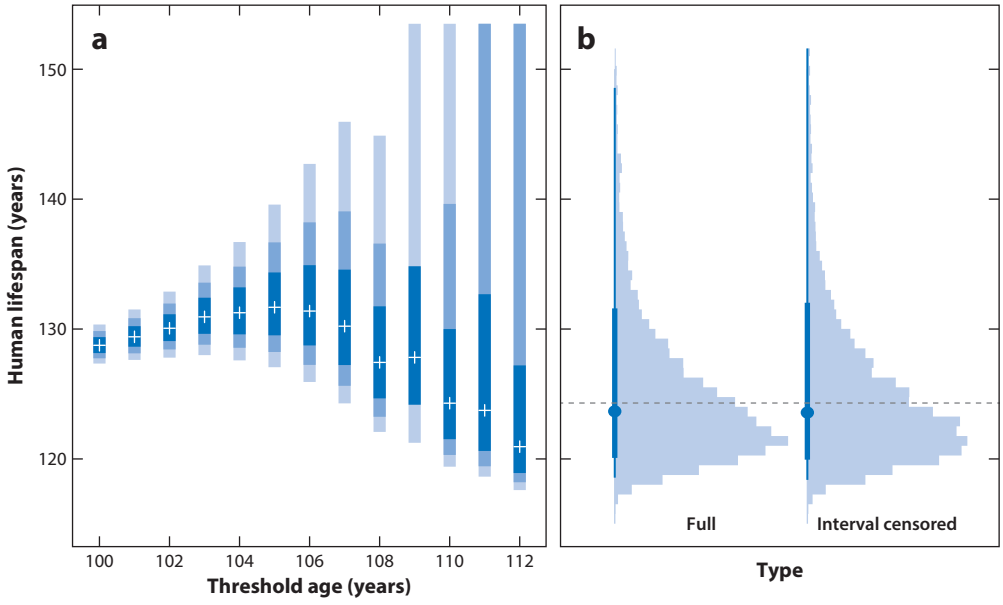

Figure 5

Analysis of Japanese data. (*a*) Maximum likelihood estimates (*crosses*) with 50%, 75%, and 90% profile likelihood confidence intervals for the endpoint ψ for a range of thresholds; some intervals end above 150 years. (*b*) Histograms of the sampling distributions for maximum likelihood estimators of ψ based on fully-observed (*left*) and interval-censored (*right*) simulated datasets. The panel shows the median (*circle*), 75%, and 95% sampling limits, and the horizontal dashed gray line is the endpoint $\psi = 124.3$ of the simulation distribution. Less than 3% of the endpoint estimates exceed 150 years, including those for the roughly 0.3% of the shape estimates that are positive.

1947 for centenarians using information from the Japanese Annual Vital Statistics Report (see **https://www.mortality.org/**). There are 10,440 semisupercentenarians in a total of 122,719 unvalidated records. Hanayama & Sibuya (2016) use a multinomial likelihood for the counts in each combination of age at death and birth cohort, based on an underlying generalized Pareto model, to compare the accumulated damage and programmed aging theories. They find better agreement with the latter, their life span estimate being 123 years using a subset of the women. Although they only consider birth cohorts for which no death has been reported for three consecutive years, up to and including 1898 for males and 1894 for females, the paper includes counts until 1898 for women, for whom $c_2 = 2014$. This selection mechanism is analogous to the hidden truncation scheme described in Section 3.2; the lifetimes are interval-censored and right-truncated.

Point estimates of the human life span do not paint the whole picture. We compute profile likelihood confidence intervals for the endpoint $\psi = u - \sigma_u/\xi$ at levels 50%, 75%, and 90%. The maximum likelihood estimates for ξ are negative for all thresholds u in the range 100, 101, . . . , 116 years, and this translates into finite point estimates $\widehat{\psi}$, but for higher thresholds the confidence intervals are strongly asymmetric, suggestive of a very large upper limit (see **Figure 5*a***). As the threshold increases, $\widehat{\psi}$ rises and then drops back, but its uncertainty increases because of the smaller sample size.

Tabulating death counts rather than publishing individual records increases data privacy but could lead to a loss of precision. To assess this, we computed the maximum likelihood estimates $\widehat{\psi}$ using both the full data and interval censoring to mimic tabulation for 10,000 datasets of size $n = 513$, with excess lifetimes simulated from the generalized Pareto model with $\sigma = 1.546$ and $\xi = -0.108$, the estimates for exceedances over $u = 110$; like the original data, the simulated

lifetimes were right-truncated. The histograms of the estimates shown in **Figure 5*b*** display downward bias, and they suggest that conventional symmetric confidence intervals centered on the estimated endpoint will be extremely poor. The sampling distributions are very similar, so the loss of information due to interval censoring is small, though it depends on the width of the age bins. This experiment suggests that this loss and any bias are immaterial for yearly bins, at least if the shape parameter is negative.

This analysis is consistent with that in Section 4.1 in suggesting that conclusions from data below around 103 years are too unstable to be useful, and it underscores the asymmetry of uncertainty for the life span ψ. Tabulation by age at death and birth cohort does not seem to increase uncertainty substantially, but it does preclude validation of individual records and therefore should be avoided.

4.3. England and Wales

In this section we illustrate the complications that truncation causes for Q–Q plotting and the radical reduction in uncertainty when extremal models are restricted. We use data on semisupercentenarians and supercentenarians in England and Wales who died between 2000 and 2014 (Off. Natl. Stat. 2016), provided by the UK Office for National Statistics. All the lifetimes are interval-truncated, and the maximum age achieved is just short of 115 years. All records for men and for lifetimes above 109 years were validated, but only a stratified sample of the remaining records was checked. There are 3,624 females and 318 males. Recent estimates (Off. Natl. Stat. 2020) conclude that the numbers of male and female semisupercentenarians have respectively increased by about 100% and 50% over the past decade, so the imbalance is diminishing.

We fitted Gompertz and generalized Pareto models for a range of thresholds. Likelihood ratio tests show no significant gender effects above 106 years. The generalized Pareto fits show strong evidence of negative shape parameters for thresholds 105–107, and both these and the Gompertz fits are consistent with plateauing at 109 years (see **Table 1**). The Q–Q plot in **Figure 6*a*** compares these data with the exponential model for threshold 109 years; the effect of the truncation is that the plot is nonmonotone. Lifetimes between 110 and 111 years fall somewhat outside the pointwise 90% confidence intervals.

Adoption of an oversimplified model may lead to underestimation of the model uncertainty. To illustrate this, we performed a Bayesian analysis, using the priors and sampling approach described at the end of Section 3.4.1 to sample from the posterior densities for the generalized Pareto, exponential, and Gompertz models. We computed 50% intervals using the half-depth method (Kay 2020) and show them in **Figure 6*b***. As the exponential hazard is constant, the intervals are very narrow. The median Gompertz and generalized Pareto hazards functions are increasing, but uncertainty for the latter encompasses constant or decreasing hazard; the latter is impossible for the Gompertz model, which is less uncertain overall.

4.4. Italy and France

ISTAT has recently produced a validated database with 3,373 women and 463 men in Italy who were at least 105 years of age at some point from January 1, 2009, to December 31, 2015; it covers birth cohorts from 1896 to 1910, and 953 individuals were alive at the end of 2015. The data, which are left-truncated and right-censored, were analyzed by Barbi et al. (2018), who fitted a Gompertz model and found a small but statistically significant cohort effect, a small and statistically insignificant effect of gender, and a plateauing of the hazard function after age 105, corresponding to an exponential excess lifetime distribution.

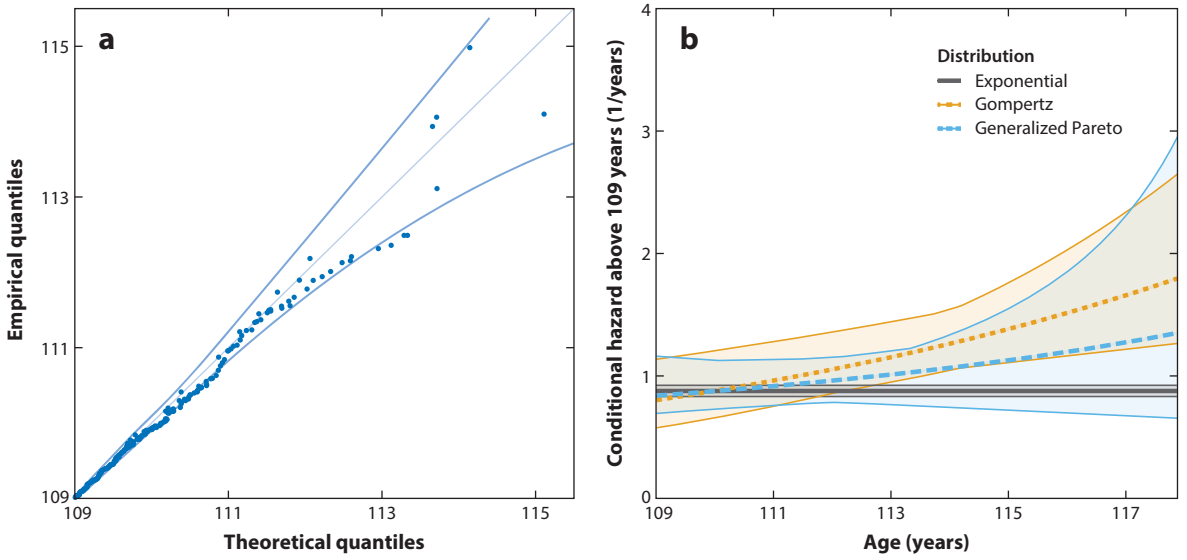

Figure 6

Analysis of exceedances above age 109 for England and Wales semisupercentenarians. (*a*) Exponential quantile–quantile plot accounting for truncation, with approximate 90% pointwise confidence intervals constructed as described in Section 3.1.4 and the **Supplemental Appendix**. (*b*) Bayesian analysis for Gompertz, generalized Pareto, and exponential hazard functions using maximal data information priors. The plot shows posterior median (*lines*) and 50% posterior predictive (*shaded areas*) intervals for the hazards.

Supplemental Material >

New French data added to the IDL in 2019 consist of 9,612 semisupercentenarians and 241 supercentenarians; all of the latter but only some of the former were validated. These data are doubly interval-truncated. Analysis of the Italian data and of most of these new French data using a variety of models led Belzile et al. (2021) to conclude that after age 108, the excess lifetimes are compatible with an exponential distribution, with no evidence of birth cohort effects. The only significant gender effect they found was for the French men, whose mean estimated survival after age 108 was 0.9 years, with the corresponding value for the Italian data and the French women being 1.46 years; the 95% confidence intervals are, for the French men and women respectively, (0.7, 1.1) and (1.36, 1.56) years. The estimated one-year survival probabilities in the French data based on an exponential distribution are 0.50 for women, with 95% confidence interval (0.48, 0.53), and 0.33 for men, with 95% confidence interval (0.25, 0.41). This may stem from a real difference in survival or from some unsuspected issue with the French data, in which the age at death for the oldest woman exceeds that for the oldest man by over a decade; it seems unlikely to be due to a lack of adjustment for multiple hypothesis testing.

Belzile et al. (2021) also use these datasets to study the power of likelihood ratio tests for detecting a finite life span: Pooling the French data with ISTAT and IDL, they find that such tests would have a power of over 85% for detecting that the endpoint of a generalized Pareto distribution lies below 130 years and of over 50% for detecting that it lies below 140 years.

4.5. International Database on Longevity

The analyses by Gampe (2010) and Rootzén & Zholud (2017) of the IDL led to the conclusion that survival above age 110 was the same for women and men, for different countries, and for persons born earlier or later, and that these lifetimes could be described by an exponential

distribution. The 2021 version of the IDL (Jdanov et al. 2021) includes about twice as many supercentenarians as the 2016 version, and the countries included have changed somewhat. Section 5 of the **Supplemental Appendix** contains updated versions of tables 1–5 of Rootzén & Zholud (2017). The general conclusions are the same: There are no detectable differences in survival between women and men or between countries. If the data are split by deaths in earlier and later periods, there is some evidence of regional changes in survival, albeit with no consistent pattern. Taken overall, there is a suggestion that the hazard of dying aged over 110 may increase from the early 2000s onwards. Such a change would be remarkable, but this apparent effect could be due to a lack of balance across regions, to possible changes in validation procedures, or simply to multiple testing, for which we have made no adjustment. If we neglect the possibility of such changes, the mean survival time for a single exponential model fitted to the new data is estimated to be 1.38 years, with 95% confidence interval (1.29,1.48), and this estimate lies close to the center of the previous 95% confidence interval. Under this model the new estimated probability of surviving one more year is 0.48, with 95% confidence interval (0.46, 0.51).

Supplemental Material >

Belzile et al. (2021) and Sections 4.1–4.4 study data on semisupercentenarians and supercentenarians from Italy, France, the Netherlands, Japan, and England and Wales. For each of these countries the analysis indicates that the hazard function appears to have plateaued by age 109, but the large uncertainty does not rule out a decreasing hazard beyond this point.

The 2021 IDL includes both semisupercentenarians and supercentenarians for Austria, Belgium, Quebec, Denmark, Germany, England and Wales, and France. **Table 6** in the **Supplemental Appendix** indicates that plateauing has also occurred well before age 109 for these countries, for which the estimated yearly probability of survival after age 109 is 0.45, with 95% confidence interval (0.36, 0.53).

5. REVIEW OF ANALYSIS STRATEGIES

Although writing $t_F = \sup\{t : F(t) < 1\} = \infty$ is natural for mathematicians, a lifetime distribution with unbounded support typically meets with skepticism. Infinity evokes the echo of immortality, but every observed lifetime has been and always will be finite, so careful translation of mathematical truths into everyday language is required. Unbounded support does not imply infinitely long lifetimes (immortality), and the fact that a certain age $\check{t}$ is highly unlikely to be surpassed under a fitted model does not imply that $t_F \leq \check{t} < \infty$.

Past strategies to assess whether $t_F < \infty$, corresponding to a cap, have been direct or indirect. In this section we briefly review some of the more prominent of them.

5.1. Tracing the Record

A superficially natural line of reasoning to investigate limits to lifetimes is to track the longest recorded lifetimes, either globally or by cohort and sometimes also by calendar year. Prompted by an early postulate that "the length of life is fixed" since "there has been no detectable change in the number of people living longer than 100 years or in the maximum age of persons dying in a given year" (Fries 1980, p. 130), record lifetimes were consulted for analysis. Both claims have long been disproved by newer and better data (Robine & Caselli 2005, Wilmoth et al. 2000), but maximum recorded lifetimes have remained an initial and also often the final point in subsequent reasoning. Argument from the intervals between records is problematic both because the population size is not stable and because these intervals are highly variable even in a stationary setting.

Fitting trend lines to annual maximum lifetimes likewise says little to nothing about t_F, especially if the size of the population from which the maximum is derived is not taken into account, and it says even less if the observation scheme is ignored. The paramount example of

this approach (Dong et al. 2016) has been shown to be flawed (e.g., Rootzén & Zholud 2017, Keiding 2018, Jdanov et al. 2021), despite a positive review by *Nature*.

Analyzing trends in maxima derived from subsets of observations and attempting inference from this sample of maxima is reminiscent of the block maximum method of extreme value theory, though it does not use this fully. Block maxima must be handled with care because of the increasing numbers of supercentenarians and the truncation and censoring of the underlying lifetimes; threshold exceedance approaches (Section 3.4.1) are preferable.

5.2. Hazard Trajectories at High Ages

Another indirect line of reasoning involves hazard trajectories at the highest ages. Barbi et al. (2018) derive a plateau of constant mortality after age 105 for the data discussed in Section 4.5, implying an exponential distribution, which has unbounded support, "underwriting doubt that any limit is yet in view" (Barbi et al. 2018, p. 1641). A constant hazard function excludes finite t_F, since for bounded support the hazard would have to increase to infinity sufficiently quickly (Section 3.1.1). A continuously increasing hazard such as that of the Gompertz model has been interpreted as indicating a finite life span (Gavrilova & Gavrilov 2020). This is incorrect: The Gompertz, Gompertz–Makeham, and many other distributions with hazards increasing to infinity have unbounded support. Finding that the Gompertz hazard holds for advanced ages would be a notable result, but even if true it would not imply that t_F is finite.

5.3. Extreme Value Analysis

Extreme value statistics, and specifically the modeling of threshold exceedances, directly addresses the question of an upper limit to the lifetime distribution. Section 3.4 presents background theory, and Section 4 gives some empirical results and discusses practical intricacies. An early study of this type was that of Aarssen & de Haan (1994), and more recently Gbari et al. (2017), Hanayama & Sibuya (2016), and Einmahl et al. (2019) have employed this strategy. The data of the last two papers were reanalyzed in Section 4. Gbari et al. (2017) fit generalized Pareto distributions to lifetimes from the Belgian national population register for the 1886–1904 birth cohorts. The study includes only extinct cohorts, so the considerations of Section 3.2 apply. They discuss different methods for threshold choice and decide on thresholds of approximately 99 years for men and 101 years for women. As a result, they get negative estimated shape parameters and finite and reasonable estimates for the maximum life span. Q–Q plots indicate a good fit for females, and perhaps less so for males, but estimates are largely determined by the bulk of the data, at ages 105 and lower. The distorting effect of grouping the observed lifetimes by year of death, as in Einmahl et al. (2019), is addressed in Section 3.2.3. In periods of increasing numbers of oldest-old, as currently seen in many countries, the uneven influx of observations at the young end of the samples biases the lifetime distribution toward lower ages.

6. CONCLUSIONS

Many, perhaps most, statistical studies of extreme human lifetimes are problematic.

A first class of problems concerns the data, which are validated with insufficient care or sometimes not validated at all. Life span estimation is particularly vulnerable to poor data, so punctilious verification of individual extreme lifetimes is essential. Another issue is the combination of datasets that bear on different phenomena and thus should not be analyzed together, and yet another is the difficulty in replicating earlier studies, owing to lack of access to data,

REPRODUCIBILITY

The R package `longevity` implements the methodology used in the article and is available for download from **https://github.com/lbelzile/longevity**. The R code used to produce the analyses is available from GitHub (**https://github.com/lbelzile/arsia-longevity**).

`LATool` is a MATLAB toolbox for life length analysis that makes alternative analyses possible. It is available on GitHub (**https://github.com/OGCJN/Is-there-a-cap-on-longevity.git**).

Data from the IDL are freely available after registration. The Dutch and Japanese data can be found from their respective publications and in `longevity`. For access to the Italian data, see Barbi et al. (2018).

apparently arbitrary or ill-documented decisions about data-handling, or a failure to provide the code to reproduce an analysis (see the sidebar titled Reproducibility).

A second class of problems is that the observation scheme is unknown or not taken into account. Data culled from news reports, social media, and the like are opportunity samples, so may be highly biased and cannot form the basis of reliable inference, whereas failure to correct for truncation and censoring, present in most observation schemes, will lead to biased estimation of the human lifetime distribution.

A third class of problems concerns extrapolation beyond the largest lifetimes yet observed to the human life span. The force of mortality increases up to age 100 or more, and there has been confusion over whether this increase implies that there is a finite human life span; this may be inferred only if the force of mortality rises indefinitely before some finite time. The natural stochastic framework for modeling rare events is extreme value theory, which involves extrapolation from asymptotic models fitted to data at finite levels. Articles using extreme value theory have tended to include data for which the limiting models are inapplicable, in some cases despite an apparently good fit due to ignoring the observation scheme. Extreme value fits to lifetimes below 105 years or even more are unstable, suggesting that the asymptotic regime has not yet been reached, and it seems plausible that only lifetimes exceeding a threshold of 109 years or so lie in the asymptotic domain. This dramatically reduces the size of the relevant sample, which roughly halves for each one-year increase in the threshold in this range.

The statistical analysis of lifetime data is full of traps for the unwary, and the widespread and natural interest in the limits to human life has led many investigators to tumble unwittingly into them. We hope that this review will lead to better studies than some past ones, even those published in supposedly highly prestigious journals.

Is there a cap on longevity? If there is one, it appears not to depend on gender or nationality, as we saw in Section 4.5. An exponential model for excess lifetimes above age 109 is consistent with the data, and neither it nor the Gompertz model allows a finite human life span.

Figure 7 summarizes evidence on the human life span ψ from the datasets considered in this article. Each panel shows the profile log-likelihood for ψ from fits of the generalized Pareto distribution to lifetimes exceeding 105, 108, and 110 years, taking truncation and/or censoring into account. The data for the IDL, England and Wales, France, and Italy were validated, but not those for Japan and the Netherlands. Lifetimes for individuals in each dataset were treated as independent and identically distributed, but despite the common upper limit, the generalized Pareto fits may differ because the other parameter of the distribution need not be the same for each dataset. In each panel the support of the fitted generalized Pareto model lies above the oldest lifetime for that dataset, so the combined support lies above the longest validated lifetime of 122 years and

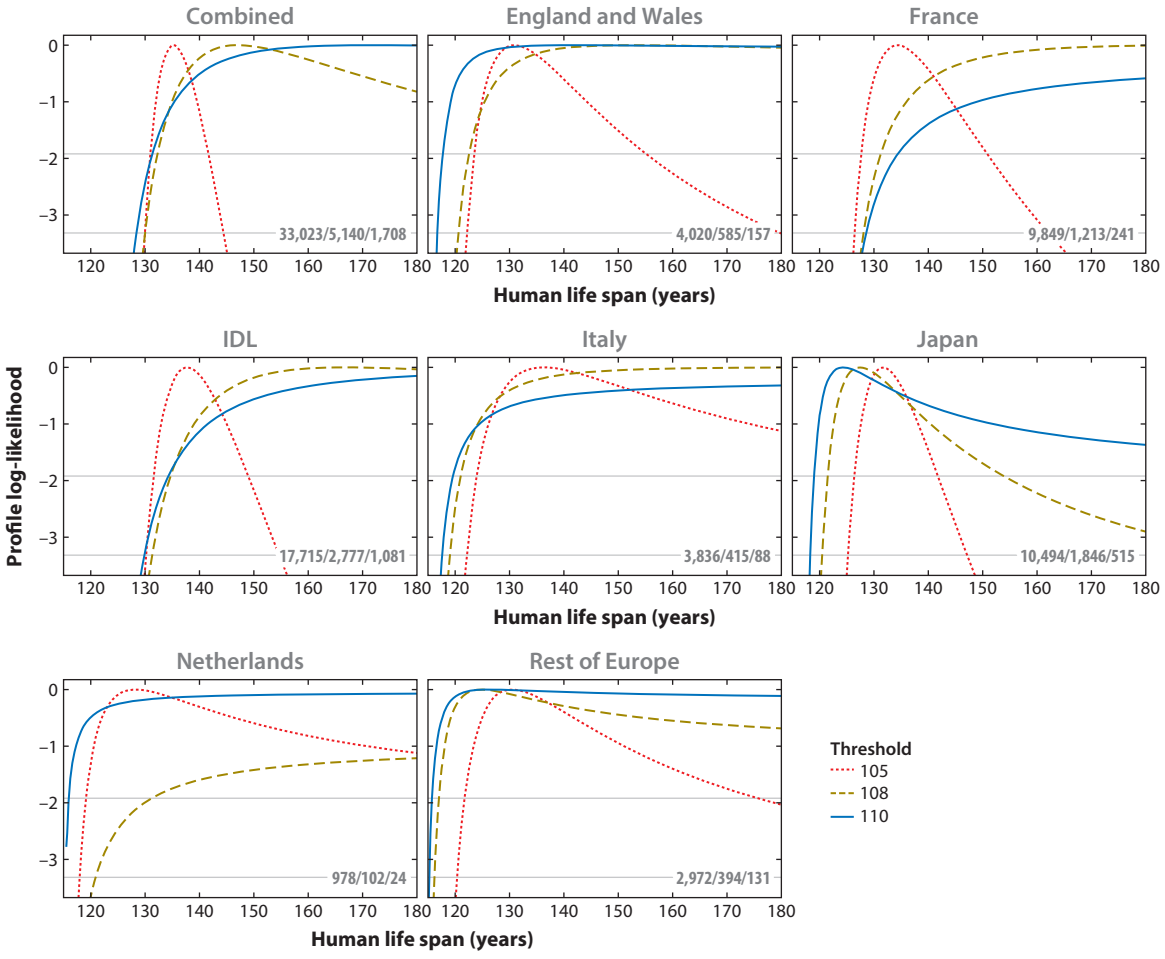

Figure 7

Profile log-likelihoods for the human life span (years) based on generalized Pareto distributions fitted to exceedances over thresholds 105 (*dotted lines*), 108 (*dashed lines*), and 110 years (*solid lines*), combined and separately for national datasets. England and Wales includes all supercentenarians from IDL and the new UK Office for National Statistics data; Italy is based on ISTAT data alone; the rest of Europe consists of IDL data for Austria, Belgium, Denmark, Finland, Germany, Norway, Spain, Sweden and Switzerland; and IDL includes the rest of Europe, England and Wales, France, and IDL Italy records, plus data from Quebec and the United States (excluding ages 105–109 for the latter). The combined fit includes Italy (i.e., ISTAT), IDL, the Netherlands, and Japan. The horizontal lines indicate limits of 95% and 99% confidence intervals based on a χ^2_1 distribution. The counts at the bottom right of each panel give the numbers of exceedances over the three thresholds. Abbreviations: IDL, International Database on Longevity; ISTAT, Italian National Statistics Institute.

164 days, or 44,724 days, for the Frenchwoman Jeanne Calment. No individuals appear twice in the combined panel.

For a threshold of 105 years—i.e., for semisupercentenarians and supercentenarians—the estimated life spans for the individual datasets lie in the range 125–140, and the 95% confidence intervals have finite upper limits except for Italy and the Netherlands. The combined life span estimate is around 135 years, with a 95% confidence interval of around (131, 143) years. For a threshold of 108 years, only Japan has an upper limit of less than 180 years for its 95% confidence interval. The other datasets put no limit to the human life span, and the combined estimate is

around 147 years, with an approximate 95% confidence interval of (133, 180+). For a threshold of 110 years, the only datasets appearing to give a 95% confidence interval with a finite upper limit are Japan and, possibly, the rest of Europe. The combined estimate seems to be above 180 years, if it is finite; the 95% confidence interval has a lower bound of 131 years or so and an apparently infinite upper bound. These confidence limits presuppose that standard likelihood theory is applicable, but even if this is not the case, the profiles suggest that the human life span lies well beyond any individual lifetime yet observed or that could be observed in the absence of major medical advances.

Gampe (2010) and Rootzén & Zholud (2017) suggested that supercentenarian lifetimes can be regarded as exponentially distributed, and the good fit of their model is confirmed by the above analysis of substantially larger datasets. Although other probability models are also compatible with these data, the exponential has a convenient interpretation due to its memorylessness: The probability that a supercentenarian will survive an extra year does not depend on his, or more usually her, current age. This probability is estimated to be 0.50, with 95% confidence interval (0.46, 0.53), so one can view this event as a coin toss. Even if the exponential model imposes no upper bound on lifetimes, it does not imply that the highest human lifetime will be very large: The probability that a particular 110-year-old will live a further 20 years is that of 20 successive heads when tossing a fair coin, a near one-in-a-million event. The anticipated increase in the number of supercentenarians will make it possible to observe a maximum recorded age at death of 130 during the present century, but a record much above this will remain highly unlikely (Pearce & Raftery 2021).

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We are grateful to Ngaire Coombs and the UK Office for National Statistics for providing updated data for England and Wales, to the International Database on Longevity staff and volunteers, to the Swiss National Science Foundation for financial support, to Tatiana Moavensadeh-Ghasnavi for helpful insights, and to Niels Keiding for very helpful comments on a draft.

LITERATURE CITED

Aalen OO, Borgan Ø, Gjessing HK. 2008. *Survival and Event History Analysis: A Process Point of View*. New York: Springer-Verlag

Aarssen K, de Haan L. 1994. On the maximal life span of humans. *Math. Popul. Stud.* 4:259–81

Andersen PK, Borgan Ø, Gill RD, Keiding N. 1993. *Statistical Models Based on Counting Processes*. New York: Springer-Verlag

Bader B, Yan J, Zhang X. 2018. Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with adjustment for false positive rate. *Ann. Appl. Stat.* 12:310–29

Barbi E, Lagona F, Marsili M, Vaupel JW, Wachter KW. 2018. The plateau of human mortality: demography of longevity pioneers. *Science* 360(6396):1459–61

Beard RE. 1971. Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. In *Biological Aspects of Demography*, ed. W Brass, pp. 57–68. London: Taylor & Francis

Bebbington M, Lai CD, Zitikis R. 2011. Modelling deceleration in senescent mortality. *Math. Popul. Stud.* 18:18–37

Beirlant J, Goegebeur Y, Teugels J, Segers JJ. 2004. *Statistics of Extremes: Theory and Applications*. New York: Wiley

Belzile LR, Davison AC, Rootzén H, Zholud D. 2021. Human mortality at extreme age. *R. Soc. Open Sci*. 8:202097

Burger O, Baudisch A, Vaupel JW. 2012. Human mortality improvement in evolutionary context. *PNAS* 109(44):18210–14

Coles SG. 2001. *An Introduction to Statistical Modeling of Extreme Values*. New York: Springer

Cournil A, Robine JM, Maier H, Gampe J, Vaupel J. 2010. The International Database on Longevity: structure and contents. In *Supercentenarians*, ed. H Maier, J Gampe, B Jeune, JW Vaupel, JM Robine, pp. 31–40. New York: Springer

Cox DR, Oakes D. 1984. *Analysis of Survival Data*. London: Chapman & Hall

Davison AC. 2018. 'The life of man, solitary, poore, nasty, brutish, and short': discussion of the paper by Rootzén and Zholud. *Extremes* 21(3):365–72

Davison AC, Smith RL. 1990. Models for exceedances over high thresholds (with discussion). *J. R. Stat. Soc. Ser. B* 52(3):393–442

de Haan L, Ferreira A. 2006. *Extreme Value Theory: An Introduction*. New York: Springer

Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. R. Stat. Soc. Ser. B* 39:1–38

Dong X, Milholland B, Vijg J. 2016. Evidence for a limit to human lifespan. *Nature* 538:257–59

Einmahl JJ, Einmahl JHJ, de Haan L. 2019. Limits to human life span through extreme value theory. *J. Am. Stat. Assoc*. 114(527):1075–80

Embrechts P, Klüppelberg C, Mikosch T. 1997. *Modelling Extremal Events for Insurance and Finance*. New York: Springer

Fries JF. 1980. Aging, natural death, and the compression of morbidity. *N. Engl. J. Med*. 303(3):130–35

Gampe J. 2010. Human mortality beyond age 110. In *Supercentenarians*, ed. H Maier, J Gampe, B Jeune, JW Vaupel, JM Robine, pp. 219–30. New York: Springer

Gavrilova NS, Gavrilov LA. 2020. Are we approaching a biological limit to human longevity? *J. Gerontol. Ser. A* 75:1061–67

Gbari S, Poulain M, Dal L, Denuit M. 2017. Extreme value analysis of mortality at the oldest ages: a case study based on individual ages at death. *N. Am. Actuar. J*. 21:397–416

Gompertz B. 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. Lond*. 115:513–85

Graunt J. 1662. *Natural and Political Observations Made upon the Bills of Mortality*. London: Thomas Roycroft

Hanayama N, Sibuya M. 2016. Estimating the upper limit of lifetime probability distribution, based on data of Japanese centenarians. *J. Gerontol. Ser. A* 71(8):1014–21

Jdanov DA, Shkolnikov VM, Gellers-Barkmann S. 2021. The International Database on Longevity: data resource profile. In *Exceptional Lifespans*, ed. H Maier, B Jeune, JW Vaupel, pp. 13–25. New York: Springer

Kaplan EL, Meier P. 1958. Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc*. 53:457–81

Kay M. 2020. `ggdist`: visualizations of distributions and uncertainty. R Package, version 2.4.0. **https://CRAN.R-project.org/package=ggdist**

Keiding N. 1990. Statistical inference in the Lexis diagram. *Philos. Trans. R. Soc. Lond. Ser. A* 332:487–509

Keiding N. 2018. Comments to Rootzén & Zholud: Human life is unlimited—but short. *Extremes* 21(2):283–86

Kestenbaum B, Ferguson BR. 2010. Supercentenarians in the United States. In *Supercentenarians*, ed. H Maier, J Gampe, B Jeune, JW Vaupel, JM Robine, pp. 219–30. New York: Springer

Makeham WM. 1860. On the law of mortality and the construction of annuity tables. *J. Inst. Actuar. Assur. Mag*. 8:301–10

Moala FA, Dey S. 2018. Objective and subjective prior distributions for the Gompertz distribution. *An. Acad. Bras. Ciênc*. 90:2643–61

Moore DF. 2016. *Applied Survival Analysis Using R*. New York: Springer

Northrop PJ. 2021. `rust`: ratio-of-uniforms simulation with transformation. *R Package*, version 1.3.12. **https://CRAN.R-project.org/package=rust**

Northrop PJ, Attalides N. 2016. Posterior propriety in Bayesian extreme value analyses using reference priors. *Stat. Sin*. 26(2):721–43
Northrop PJ, Coleman CL. 2014. Improved diagnostic plots for extreme value analyses. *Extremes* 17:289–303
Off. Natl. Stat. 2016. *Accuracy of official high-age population estimates, in England and Wales: an evaluation*. Tech. Rep., Off. Natl. Stat., London
Off. Natl. Stat. 2020. *Estimates of the very old, including centenarians, UK: 2002 to 2019*. Tech. Rep., Off. Natl. Stat., London
Papastathopoulos I, Tawn JA. 2013. Extended generalised Pareto models for tail estimation. *J. Stat. Plan. Inference* 143(1):131–43
Pearce M, Raftery AE. 2021. Probabilistic forecasting of human maximum lifespan by 2100 using Bayesian population projections. *Demographic Res*. 44:1271–94
Perks W. 1932. On some experiments in the graduation of mortality statistics. *J. Inst. Actuar*. 61(1):12–57
Pickands III J. 1975. Statistical inference using extreme order statistics. *Ann. Stat*. 3(1):119–31
Poulain M. 2010. On the age validation of supercentenarians. In *Supercentenarians*, ed. H Maier, J Gampe, B Jeune, JW Vaupel, JM Robine, pp. 3–30. New York: Springer
Resnick SI. 2006. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. New York: Springer
Robine JM, Caselli G. 2005. An unprecedented increase in the number of centenarians. *Genus* 61(1):57–82
Rootzén H, Zholud D. 2017. Human life is unlimited—but short (with discussion). *Extremes* 20(4):713–28
Scarrott C, MacDonald A. 2012. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT* 10:33–60
Smith RL. 1987. *Approximations in extreme value theory*. Tech. Rep. 205, Cent. Stoch. Proc., Univ. N. C., Chapel Hill
Thatcher AR. 1999. The long-term pattern of adult mortality and the highest attained age (with discussion). *J. R. Stat. Soc. Ser. A* 162:5–43
Thatcher AR, Kannisto V, Vaupel JW. 1998. *The Force of Mortality at Ages 80 to 120*. Odense, Den.: Odense Univ. Press
Therneau TM, Grambsch PM. 2000. *Modeling Survival Data: Extending the Cox Model*. New York: Springer-Verlag
Tsai WY, Jewell NP, Wang MC. 1987. A note on the product-limit estimator under right censoring and left truncation. *Biometrika* 74(4):883–86
Turnbull BW. 1976. The empirical distribution function with arbitrarily grouped, censored and truncated data. *J. R. Stat. Soc. Ser. B* 38:290–95
Vaupel JW, Manton KG, Stallard E. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16(3):439–54
Wadsworth JL. 2016. Exploiting structure of maximum likelihood estimators for extreme value threshold selection. *Technometrics* 58:116–26
Wakefield JC, Gelfand AE, Smith AFM. 1991. Efficient generation of random variates via the ratio-of-uniforms method. *Stat. Comput*. 1:129–33
Waller LA, Turnbull BW. 1992. Probability plotting with censored data. *Am. Stat*. 46(1):5–12
Wilmoth JR, Deegan LJ, Lundstrom H, Horiuch S. 2000. Increase of maximum life-span in Sweden, 1861–1999. *Science* 289:2366–68
Zellner A. 1977. Maximal data information prior distributions. In *New Developments in the Applications of Bayesian Methods*, ed. A Aykac, C Brumat, pp. 211–32. Amsterdam: North-Holland