The World Trend in Maximum Life Span

JOHN R. WILMOTH
JEAN-MARIE ROBINE

The maximum observed life span is known to have been rising slowly for at least 140 years in Sweden and has presumably been rising in other industrialized countries as well (Wilmoth et al. 2000; Wilmoth and Lundström 1996). If we consider men and women together, the Swedish maximum age at death in a calendar year rose from around 102 years in the 1860s to around 109 in the 1990s. This result is significant because no other country offers such a long series of high-quality data that can be used to study maximum life span. An obvious question is whether the Swedish trend is typical of the world, or at least some part of the world.

Only four Swedes are known to have achieved the age of 110 years. In the world as a whole, however, the number of "validated super-centenarians," or individuals aged 110 or older for whom adequate documentation is available, is increasing (Robine and Vaupel 2001). Several of these cases lie well above the Swedish trend, and the difference is large enough to require a rigorous explanation. Perhaps in some important way Sweden is not representative of the rest of the world, or even of the industrialized world. Or perhaps the world records are higher than the Swedish ones merely because they are drawn from a larger population. If the latter explanation (or some variant of it) is true, then it should be possible to speculate about the world trend in maximum life span by extrapolation from the Swedish trend.

Empirical evidence regarding trends in the human life span

It is well known that age is often misreported in human populations, especially at older ages. Furthermore, it is often impossible to confirm or disprove the accuracy of a given claim absent documentary evidence. Today's centenarians and super-centenarians were born, obviously, more than 100 or 110 years ago, when record keeping was less complete than it is today. In only a handful of populations are complete birth records available from the late nineteenth century that would permit an exhaustive check of all cases of extreme longevity reported today. A complete system of Swedish national statistics dating from 1749 helps to explain the accuracy of death records at extreme ages beginning in the 1860s. For countries with less complete records, it may nevertheless be possible to certify the accuracy of an individual case of extreme longevity by assembling the available evidence, carefully checking its consistency, and, whenever possible, interviewing the person in question as a final check on the coherence of the case (Jeune and Vaupel 1999).

Keeping these remarks about data quality in mind, we begin with a review of the empirical evidence. The unmistakable conclusion is that the maximum observed life span has been increasing for decades and possibly for one or more centuries, not only in Sweden but also for the human population as a whole. We must first resolve, however, whether the slope of the increase is similar in different populations. For this purpose, we review the Swedish evidence since 1861, analyze individual records of confirmed super-centenarians, and study trends in the maximum age at death for seven countries that have complete and validated records since at least 1950.

Sweden

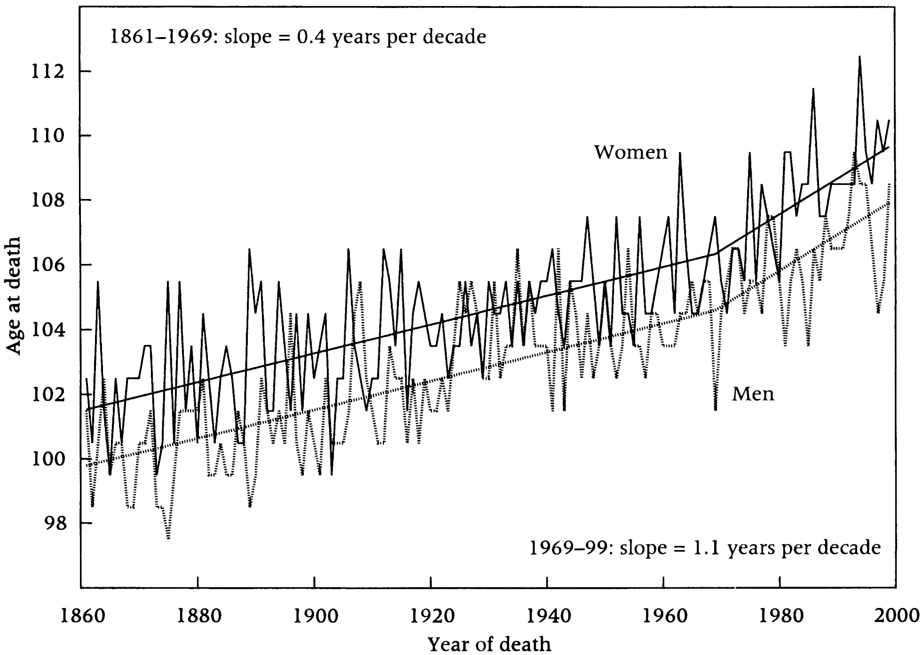
The increase in the maximum age at death by sex in Sweden is shown in Figure 1. This graph, published in Wilmoth et al. (2000), serves as a reminder of current knowledge about trends in maximum human life span. The slope of the parallel regression lines (expressed in years per decade) is 0.4 before 1969 and 1.1 afterward. The major components of this increase can be studied using a statistical model, as reviewed below.

Individual cases

The increasing number of super-centenarians in the world has only recently been documented (Robine and Vaupel 2001). Figure 2A shows the maximum age at death for calendar years since 1959 based on Robine and Vaupel's published list of validated super-centenarians (thus, by definition the distribution is truncated below age 110). For every year since 1977, the oldest validated death in the world was at age 110 or above. Given this information, it appears that maximum age at death in the world increased at a pace of 3.1 years per decade during 1977–2000, thus much more quickly than the Swedish trend. We must, however, consider two sources of bias.

First, it is possible that some of these "validated" cases are not genuine. Robine and Vaupel (2001) distinguish different levels of validation and propose that further analyses should be based on "3-star validated" cases,

FIGURE 1 Maximum age at death as annually observed, by sex, and trend lines, Sweden 1861–1999



SOURCE: Wilmoth et al. (2000).

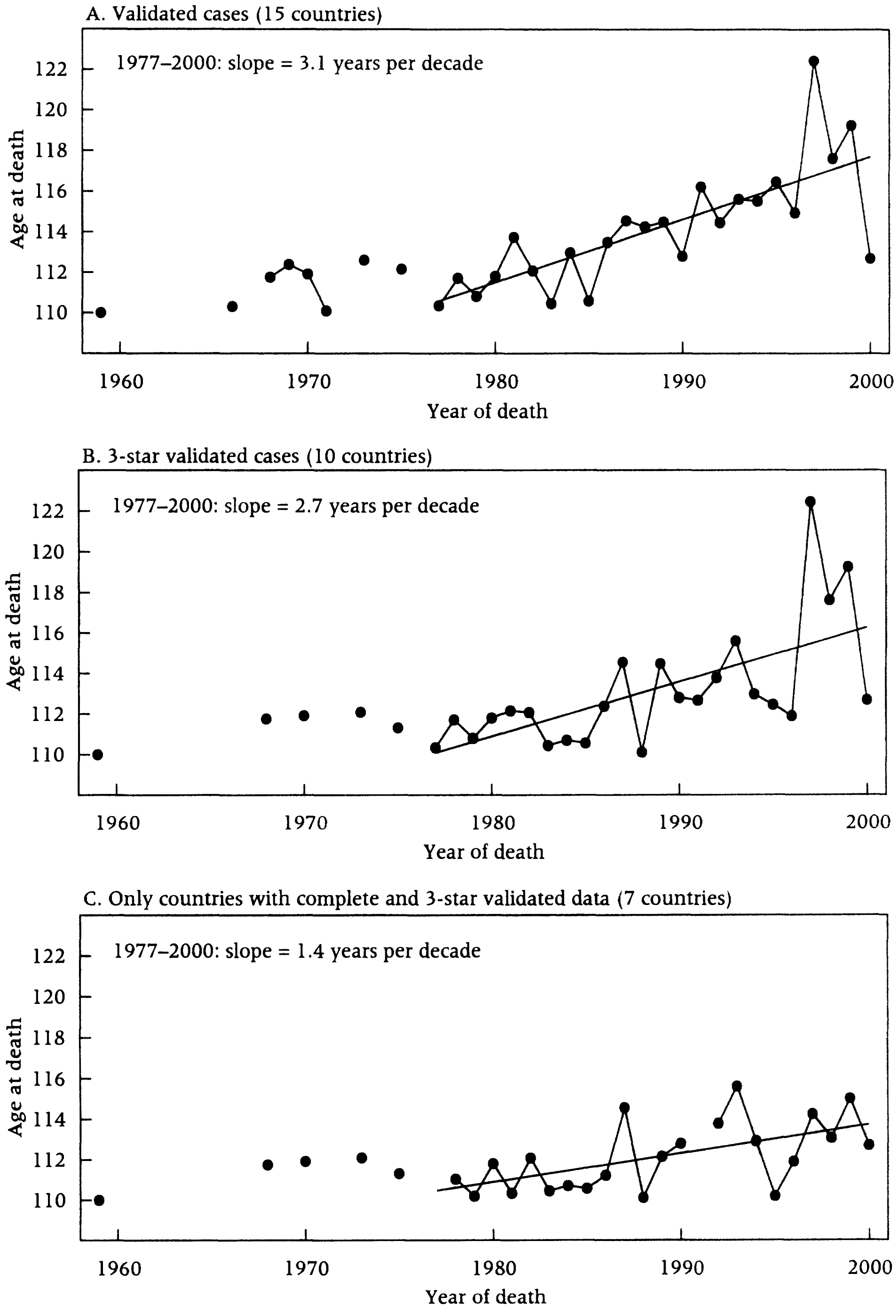
indicating that birth and death records (or photocopies thereof) have been brought together and compared side by side to verify the reported age. Figure 2B shows that when we limit the analysis to such cases, the slope of increase is slightly less, or about 2.7 years per decade during 1977–2000.

Other countries with complete and valid data

Second, it is possible that earlier cases of super-centenarians have escaped notice, introducing another source of upward bias to the slope shown in Figures 2A and 2B. Robine and Vaupel (2001) suggest therefore that it may be appropriate to limit our attention to countries with an exhaustive system for validating apparent super-centenarians. There are at present only seven such countries: Belgium, Denmark, England and Wales,[1] Finland, the Netherlands, Norway, and Sweden. As shown in Figure 2C, the slope is reduced further if we limit ourselves to such cases. In fact, at 1.4 years per decade during 1977–2000, the rate of increase in the maximum age at death for these seven countries is only slightly higher than the value of 1.1 noted earlier for Sweden during 1969–99.

Obviously, the truncation of age at 110 years limits the time period of the analysis as well, since super-centenarians were still quite rare before the

FIGURE 2 Maximum age at death among super-centenarians

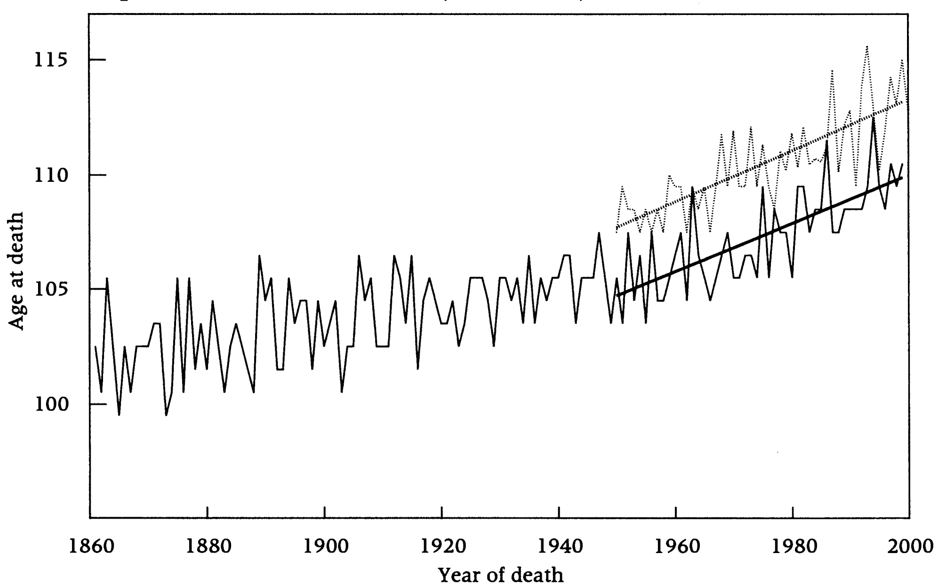


NOTES: (1) Seven countries provide the data shown in 2C: Belgium, Denmark, England and Wales, Finland, the Netherlands, Norway, and Sweden. Although data are complete only for these seven, three additional countries contributed at least some cases of "3-star validated" super-centenarians used to construct 2B: Canada, France, and the United States. Using a less rigorous selection criterion, six more countries contributed cases of "validated" super-centenarians shown in 2A: Australia, Italy, Japan, Scotland, South Africa, and Spain. Thus, since England, Wales, and Scotland are all part of the United Kingdom, a total of 15 countries have contributed data to this analysis. (2) Trend lines for Figure 2 were fit by ordinary least squares. (3) In 2C, missing values for 1977 and 1991 were omitted from the regression.
SOURCE: Robine and Vaupel (2001).

1970s. For six of these seven countries (with the exception of Belgium), however, we are able to fill in the missing parts of the trend using national data. Thus, Figure 3 shows the maximum age at death, for men and women combined, in Sweden and in the group of "complete and validated" countries (omitting Belgium). For the latter group, the trend in Figure 3 is based on data from the super-centenarian database whenever possible (i.e., a death that occurred above age 110 in a given calendar year) and on national death statistics otherwise. The two trend lines, based on data for 1950–99, are separated by about 3.1 years but are very nearly parallel, each with a slope around 1.1 years per decade.

In light of this analysis, we conclude that the slope of the trend in maximum life span has been the same in Sweden and in the other "complete and validated" countries since at least 1950. Furthermore, the increase accelerated over this time period and was faster in the last quarter of the twentieth century. It remains to be seen, however, whether the observed difference in level (of about 3.1 years) can be explained simply as a function of population size, or whether other factors also affect the difference.

FIGURE 3 Maximum age at death with trend lines, Sweden 1861–1999 and "complete and validated" cases (6 countries), 1950–2000



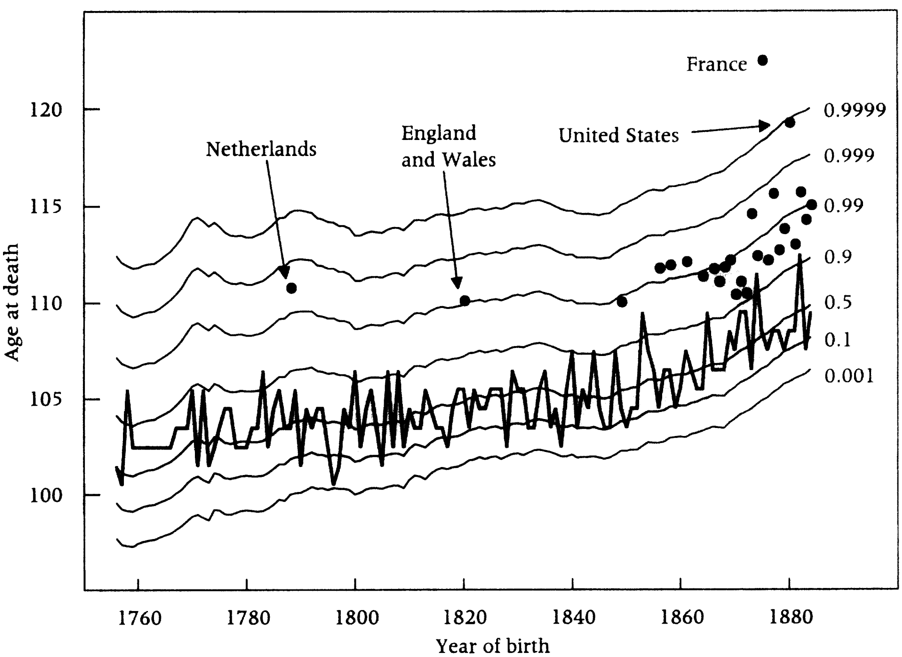
NOTES: (1) Six countries underlie the trend shown here for 1950–2000: Denmark, England and Wales, Finland, the Netherlands, Norway, and Sweden. Although Belgium contributes to the universe of "complete and validated" super-centenarians (Figure 2C), it lacks mortality data above age 100 in the Kannisto–Thatcher database, which is used here to fill in the trend below age 110. (2) Trend lines for 1950–99 were fit by ordinary least squares.
SOURCES: Wilmoth et al. (2000); Robine and Vaupel (2001); Kannisto–Thatcher Database on Old-Age Mortality.

Statistical models of extreme longevity

The maximum age at death can be modeled statistically by treating it as an extreme value of a random sample from some probability distribution (Gumbel 1937, 1958; Aarssen and de Haan 1994; Wilmoth et al. 2000). With such models, it is possible to "predict" the maximum age at death for a cohort of individuals born in the same year. Furthermore, it is possible to compute the complete distribution of the maximum age at death for a given cohort. In other words, the observed value is only one observation from a range of possible values, and the probability of observing any value in that range is represented by the said distribution. Figure 4 shows the trend in the Swedish maximum age at death by cohort, for men and women combined, along with quantiles of the estimated distribution of the extreme value. For example, there is a 50/50 chance that the maximum age at death observed in a given calendar year lies above or below the line corresponding to the median (i.e., the 0.5 quantile). This figure also shows all of the 3-star validated super-centenarians in the world. Clearly, these cases lie well above the Swedish trend.

We attempt to explain this difference using statistical models of extreme values. We discuss such models in general, describe their application

FIGURE 4 Maximum age at death in Sweden with predicted quantiles, compared to all 3-star validated super-centenarians in the world, birth cohorts of 1756–1884



to the Swedish trend in the maximum age at death, and review the principal results from an earlier study (Wilmoth et al. 2000). All of the new calculations presented here treat the Swedish example as a baseline, so it is important to understand how the earlier results were derived.

Extreme value distributions

Suppose that $S(x)$ is the probability of survival from birth to age x for an individual chosen at random from a cohort of N births.[2] Then, the probability that the maximum age at death for this cohort lies above age x is given by $S^N(x) = 1 - [1 - S(x)]^N$. Accordingly, the maximum age at death is itself a random variable with a probability distribution, and this distribution is determined by N and the $S(x)$ function or, alternatively, by N and the probability distribution of ages at death, given by the function

$$f(x) = \frac{-dS(x)}{dx}.$$

In the earlier analysis, the mortality experience of the Swedish cohort born in year t was described by a vector of age-specific death rates, $\mathbf{M_t} = (M_{0t}, M_{1t}, \ldots, M_{79t}, \hat{M}_{80t}, \ldots, \hat{M}_{119t})$, which was transformed into cohort survival probabilities, $S(x,t)$.[3]

Below age 80 years, death rates were observed values for cohort t, whereas at higher ages they were derived from a logistic model.[4] Wilmoth et al. (2000) provides further information on how this model was applied in the Swedish case. However, one detail requires special attention. The logistic model implies that death rates approach a fixed upper limit, denoted here by r. However, the level of r is unknown and difficult to determine. In the analysis of Swedish data, the upper asymptote was fixed at 1.25 because that value maximized a global measure of goodness-of-fit. Although such an assumption has no strong justification, it was convenient since cohort-specific estimates of r are highly unstable owing to random fluctuations in death rates at older ages. Moreover, sensitivity analysis confirmed that the main results of interest in the Swedish study did not vary significantly with alternative choices of r.

In effect, the value of r in the previous study was used as a kind of "tuning constant." A lower value of r would have moved the entire distribution of the maximum age at death upward, whereas a higher value would have pushed it downward. The value of 1.25 was chosen because it placed the distribution at the correct level for the observed data. Since a formal criterion was used for choosing the optimal value, the methodology used in the Swedish study offers a new means of estimating r based on extreme values rather than on the shape of the age curve of mortality at lower ages. The latter method was used, for example, by Thatcher et al. (1998), who typically found much lower values of *r*, often below 1.[5] It is also worth noting that the simple model of old-age mortality proposed by Kannisto is in fact the logistic model with an asymptote of 1.[6]

Thus, although it describes the Swedish collection of extreme values quite well, an asymptote of 1.25 in the logistic model seems impossibly high. Before attempting to extrapolate from the Swedish trend to the rest of the world, we first examine whether the risk of mortality at extremely high ages in Sweden is typical of other industrialized countries. We begin by reviewing one other key result of the earlier Swedish study, which will help to guide our further analysis.

Explaining the rise in maximum life span

What factors determine the level of the maximum age at death and its changes over time? The Swedish study addressed this question by examining the components of the increase in the median of the estimated distribution of the maximum age for each cohort. The conclusion was that 72.5 percent of the rise in the Swedish maximum age at death from 1861 to 1999 is attributable to reductions in death rates above age 70, with the rest due to increased numbers of survivors to old age (both larger birth cohorts and increased survivorship from infancy to age 70).

Even though mortality below age 70 fell precipitously over this time period, that change accounts for only 16 percent of the rise in the maximum age at death. Death rates at older ages fell less markedly, but the impact of this change on extreme values was much greater. The annual number of births in Sweden increased only slightly over the relevant range of cohorts and, for that reason, had a relatively minor impact on the trend in the maximum age at death. However, a larger change in the number of births would have had a much more significant effect, as demonstrated below. Thus, to explain the differences seen earlier in the level of the maximum age at death between populations, we must consider primarily any differences in population size and in patterns of old-age mortality.

Comparing trends in the maximum age at death

We have demonstrated that the level of the trend in the maximum age at death differs depending on the population being considered. On the other hand, in all cases where we have examined complete and validated records for a population, the slope of this trend appears to be similar. Thus, our task is to explain differences in the level of the trend between populations. As noted already, two major factors must be considered: the size of the "eligible" population, and the level and pattern of old-age mortality.

The "eligible" population

It is difficult to define the universe from which selected cases of extreme longevity have been drawn. We must consider the fact that any person who is eligible to become a validated super-centenarian must live in a country or region in which the necessary documentation is available. One strategy, therefore, is to add up the populations of all such areas and to apply the same calculations made earlier for Sweden to this larger population. Of course, it is difficult to determine such a collection of countries, since all included areas should have reliable vital statistics not only today but also more than 100 years ago. Furthermore, we would ideally need to compare not the size of the total populations, but rather the size of birth cohorts from over 100 years ago. To simplify the task (with minimal loss of accuracy), we derive ratios of relative population size from estimates of national populations in the year 2000.

For the present purpose, we consider three populations: (1) the population that yielded the "complete and validated" list of super-centenarians, (2) the population that produced the 3-star validated super-centenarians, and (3) the total population of the world with a mortality experience that could plausibly yield cases of world record longevity. The first two populations are used for a comparison between the Swedish trend in maximum life span and the data on super-centenarians presented earlier (supplemented by national data in the case of the "complete and validated" countries). The last population is used to speculate about the world trend in maximum life span.

For each of these cases, we specify a preferred scenario. In the second and third cases, we also investigate alternative scenarios. The population sizes for each scenario are based on Table 1, which gives estimates of the year 2000 population for all countries included in either the super-centenarian database or the original Kannisto–Thatcher Database on Old-Age Mortality (Kannisto 1994). The reasoning used for each scenario is described in a later section.

Mortality variation at older ages

Mortality rates vary across national populations, even among low-mortality countries. Table 2 summarizes mortality estimates derived from the Kannisto–Thatcher (KT) database for countries with reliable data covering the age range above 80 years.[7] From this table, it is clear that Sweden has an atypical mortality pattern at older ages. As summarized in Table 3, Swedish mortality is relatively low at ages 80–89 but relatively high at age 100+. This is true whether the point of comparison is an aggregate of 13 countries with complete, high-quality data (see Table 2) or our smaller collection of "complete and validated" countries (excluding Belgium in this instance because it lacks data above age 100 in the KT database). These differences

TABLE 1 Countries included in the super-centenarian and in the Kannisto–Thatcher database

Country	Super-centenarian database: Full list	3-star validated	Complete and validated	Kannisto-Thatcher database	Population in 2000 (millions)
Australia	X			X	19.2
Austria				X	8.1
Belgium	X	X	X	X	10.2
Canada	X	X			30.8
Czech Republic				X	10.3
Denmark	X	X	X	X	5.3
England and Wales	X	X	X	X	52.7
Estonia				X	1.4
Finland	X	X	X	X	5.2
France	X	X		X	59.4
Germany				X	82.1
Hungary				X	10.0
Iceland				X	0.3
Ireland				X	3.8
Italy	X			X	57.8
Japan	X			X	126.9
Latvia				X	2.4
Luxembourg				X	0.4
Netherlands	X	X	X	X	15.9
New Zealand				X	3.8
Norway	X	X	X	X	4.5
Poland				X	38.6
Portugal				X	10.0
Scotland	X			X	5.1
Singapore				X	4.0
Slovakia				X	5.4
South Africa	X				43.4
Spain	X			X	39.5
Sweden	X	X	X	X	8.9
Switzerland				X	7.1
United States	X	X			275.6
Total population in 2000 (millions)	760.4	468.5	102.7	598.5	946.3

NOTES: (1) Following Robine and Vaupel (2001), the "full list" used in this analysis includes cases of "validated" super-centenarians, whose age was confirmed using some form of *prima facie* documentary evidence. Cases designated as "3-star validated" have passed a more rigorous test, consisting of a side-by-side comparison of official birth and death records (or photocopies thereof). In the seven countries designated as "complete and validated," it has been possible to apply the 3-star criterion to all cases of alleged super-centenarians, of which only confirmed cases are analyzed here. (2) The original Kannisto–Thatcher database included 28 countries (see Note 1 at the end of this chapter). The list shown here has the same number: although the former Czechoslovakia has been split in two, East and West Germany have been combined. (3) The Kannisto–Thatcher database includes only the non-Maori population of New Zealand and the Chinese population of Singapore. However, population figures shown here refer to the full populations of both areas.
SOURCES: (1) All population estimates (except for areas within the United Kingdom) refer to mid-2000 and are from the Population Reference Bureau, *World Population Data Sheet 2001*, http://www.prb.org/pubs/wpds2000/. (2) Population estimates for parts of the United Kingdom (England and Wales, and Scotland) refer to mid-1999 and are from http://www.statistics.gov.uk/ (3) The super-centenarian database used here is described by Robine and Vaupel (2001). The Kannisto–Thatcher database is described by Kannisto (1994).

TABLE 2 Death rates by sex above age 80 for 13 countries with complete, high-quality data, circa 1960–98

	Women				
Country	**80–89**	**90–99**	**100+**	**105+**	**110+**
Austria	0.123	0.282	0.550	0.716	0.462
Denmark	0.103	0.247	0.512	0.704	0.429
England and Wales	0.107	0.238	0.460	0.643	0.684
Finland	0.117	0.265	0.500	0.687	0.429
France	0.100	0.240	0.480	0.611	0.485
Germany, West	0.116	0.269	0.514	0.576	0.921
Iceland	0.092	0.227	0.488	0.643	na
Italy	0.114	0.268	0.488	0.573	1.091
Japan	0.093	0.223	0.425	0.513	0.446
Netherlands	0.100	0.245	0.501	0.648	2.571
Norway	0.103	0.250	0.489	0.717	0.857
Sweden	0.100	0.244	0.500	0.623	0.909
Switzerland	0.099	0.245	0.504	0.650	6.000
13 countries	0.105	0.246	0.474	0.601	0.613
(95% C.I.)	(0.105, 0.105)	(0.246, 0.246)	(0.471, 0.476)	(0.589, 0.614)	(0.521, 0.705)

	Men				
Country	**80–89**	**90–99**	**100+**	**105+**	**110+**
Austria	0.158	0.326	0.601	0.857	na
Denmark	0.141	0.298	0.596	0.821	na
England and Wales	0.151	0.297	0.533	0.635	0.429
Finland	0.153	0.307	0.620	0.909	na
France	0.139	0.292	0.576	0.585	0.794
Germany, West	0.156	0.322	0.602	0.571	1.103
Iceland	0.119	0.253	0.659	6.000	na
Italy	0.146	0.311	0.560	0.645	0.857
Japan	0.131	0.273	0.499	0.663	0.783
Netherlands	0.140	0.292	0.568	0.855	na
Norway	0.137	0.295	0.532	0.516	0.375
Sweden	0.137	0.301	0.581	0.712	na
Switzerland	0.136	0.295	0.567	0.721	na
13 countries	0.143	0.298	0.554	0.638	0.729
(95% C.I.)	(0.143, 0.143)	(0.297, 0.298)	(0.548, 0.559)	(0.606, 0.671)	0.417, 1.040)

NOTES: (1) Data for all countries are available during 1960–98 (inclusive) with the following exceptions: Data for Iceland begin in 1961, data for Italy end in 1993, data for Denmark end in 1999, and data for West Germany are missing above age 100 during 1960–63 and above age 105 in 1975. (2) Confidence intervals for death rates in the 13-country aggregate population were computed by assuming Poisson variability of death counts (Brillinger 1986) and using a normal approximation. (3) Death rates above age 110 are marked as "na" in the table if there were zero person-years of exposure at these ages over the given time period.
SOURCE: Original calculations using the Kannisto–Thatcher Database on Old-Age Mortality.

TABLE 3 Ratio of death rates for two groups of countries compared to Sweden, by age and sex, circa 1960–98

	13-country aggregate			Complete and validated		
Age group	Women	Men	Total	Women	Men	Total
80–89	1.05	1.04	1.03	1.05	1.06	1.04
90–99	1.01	0.99	0.99	0.99	0.99	0.97
100+	0.95	0.95	0.94	0.95	0.95	0.94

NOTES: (1) The group of 13 countries is listed in Table 2. (2) The "complete and validated" countries are listed in Table 1. However, this comparison excludes Belgium, which lacks mortality data above age 100 in the KT database.
SOURCE: Original calculations using the Kannisto–Thatcher Database on Old-Age Mortality.

should be taken into account when modeling the trend in maximum life span for other groups of countries based on the Swedish experience.

Scenarios consistent with "validated" cases

We now define a variety of scenarios in an attempt to explain the validated cases of extreme longevity in comparison to the Swedish trend. We build on the earlier model for Sweden (see Figure 4). Our strategy is to specify scenarios in which population size and the level of old-age mortality are expressed relative to comparable values for the Swedish population. However, we define explicitly the level of r, the asymptotic upper bound on death rates at older ages. The various scenarios are summarized in Table 4.[8]

The first row of Table 4 shows the baseline values for each of the parameters in the current model. All parameters except r equal one, because they are expressed relative to the baseline (Swedish) model. The next three rows investigate the effect of changing population size alone. The effect is expressed in terms of the median increase, for the most recent 40 cohorts, in the 50th percentile of the estimated distribution of the maximum age at death. For a population 10 times as large, the median maximum age at death should be 2.86 years higher. For a population 100 or 1,000 times as large, the median should be 5.47 and 7.90 years higher, respectively. It is essential to note the diminishing importance of increasing population size. Also, we should remember that a population 1,000 times as large as that of Sweden (with 8.9 million people in 2000) would be larger than the current population of the world.

The next row of Table 4 defines a scenario for countries with "complete and validated" super-centenarian data. The population in this situation is well defined. We note that there were 102.7 million persons in these countries in 2000, almost 12 times the population size of Sweden.[9] Because the relative population size is unambiguous and because we have direct information on the relative level of mortality, there is no alternative scenario in this case. In addition to a 12-fold increase in population size, we assume a 4 percent increase in mortality at age 84 (based on the information for ages 80–89 in the

TABLE 4 Scenarios to predict the trend and distribution of the maximum age at death for four populations

	Compared to Swedish baseline for cohort *t*, relative level of:						
	Annual number of births	Probability of survival to age 80	Force of mortality at age 84	Mortality increase with age	Upper limit of force of mortality	Median difference in the median maximum age at death (in years)	
Scenario	$\frac{N(t)}{N_0(t)}$	$\frac{S(80,t)}{S_0(80,t)}$	$\frac{\mu(84,t)}{\mu_0(84,t)}$	$\frac{\theta(t)}{\theta_0(t)}$	r	Predicted	Observed
Baseline (Sweden)	1	1	1	1	1.25	0	na
Population size (1)	10	1	1	1	1.25	2.86	na
Population size (2)	100	1	1	1	1.25	5.47	na
Population size (3)	1,000	1	1	1	1.25	7.90	na
Complete and validated	12	1	1.04	1	1.18	3.01	3.00
3-star validated A	18	1	1.03	1	1.02	4.51	4.50
3-star validated B	42	1	1.06	1	1.15	4.47	4.50
World I	85	1	1.03	1	1.02	6.45	na
Size only	85	1	1	1	1.25	5.29	na
Mortality only	1	1	1.03	1	1.02	0.62	na
World II	106	1	1.06	1	1.02	6.44	na
World III	106	1	1.06	1	1.15	5.53	na

NOTES: (1) The baseline scenario, denoted by the subscript 0, depicts the evolution of the maximum age at death in Sweden for cohorts born from 1756 to 1884 (Wilmoth et al. 2000). See text for an explanation of the other scenarios. (2) The second-to-last column records the median difference (between a given scenario and the baseline) in the 0.5 quantile of the estimated probability distributions of the maximum age at death for the most recent 40 cohorts (i.e., those born 1845–84). (3) For the "complete and validated" countries (excluding Belgium), the last column gives the observed median difference (compared to Swedish data) in the actual maximum age at death for calendar years 1950–99. For countries with 3-star validated super-centenarians, the last column gives the median difference (compared to Swedish data) in the maximum age at death for all cohorts born after 1849, with missing values (for cohorts with no validated super-centenarians) set to age 109. (4) See Notes 4 and 8 at the end of this article for definitions of key mathematical symbols. (5) As noted in the text, population size is taken as a proxy for size of birth cohort.

total population of such countries, as seen in Table 3). However, the value of the asymptote, r, is chosen to obtain a good match between data and model. Thus, the value $r = 1.18$ yields a close fit. This assumption is not implausible, given that observed mortality rates above age 100 in these countries are about 6 percent lower than in Sweden (see Table 3).

Next, we attempt to reconstruct the population that yielded the 3-star validated centenarians. We include all 102.7 million persons residing in the "complete and validated" countries, plus a fraction of the populations for other countries where 3-star validated centenarians have been found. This fraction equals the number of 3-star validated super-centenarians as a proportion of the total number expected in each country (based on the prevalence of super-centenarians in the "complete and validated" countries). The expected numbers are taken from Robine and Vaupel (2001). For example, for the United States only two super-centenarians are included on this list, whereas 184 would have been expected. Therefore, only about 1 percent of the US population, or 3.0 million, is included in the population from which these super-centenarians have been drawn. Making similar calculations for Canada and France, and adding those numbers to the 102.7 million for the other seven countries, yields a population of 157.6 million, or about 18 times the size of Sweden's population.

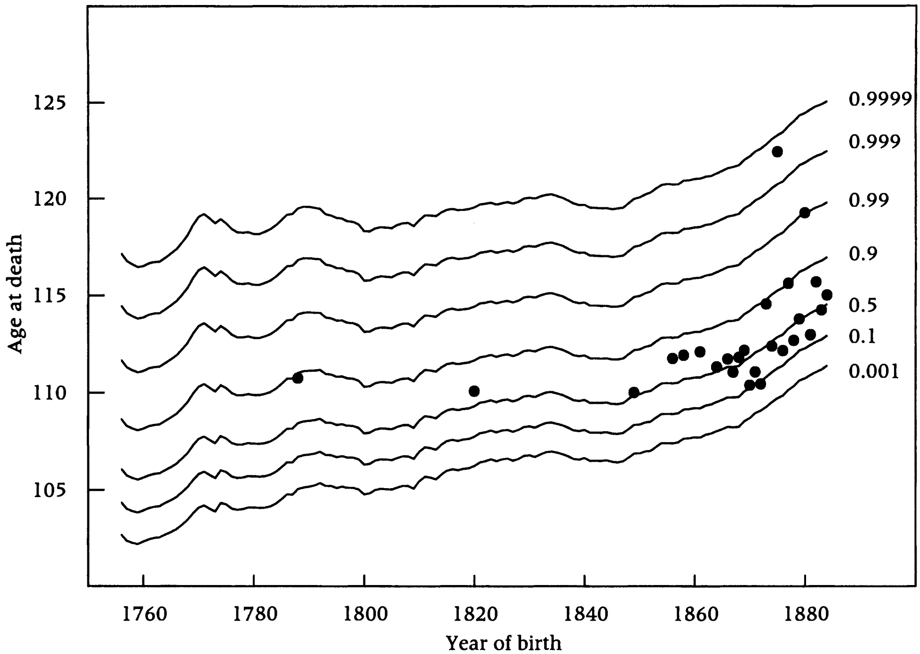
To complete the scenario, we must also specify the relative mortality level at age 84. The three countries that have been added to the list (France, Canada, and the United States) all have relatively low death rates at older ages, similar to or slightly below Sweden. However, these three countries contribute only about a third of the total population under this scenario (an 18-fold increase as opposed to a 12-fold increase in the previous case). Therefore, we set the relative mortality level at 1.03 in this scenario, only slightly lower than for the "complete and validated" countries alone. As before, the value of the asymptote, $r = 1.02$, was chosen to obtain a close match to the data. As noted earlier, however, a value of r around 1 is closer to previous estimates of this parameter by other authors (Thatcher et al. 1998).

Alternatively, one might argue that all of the remaining populations in the Kannisto–Thatcher database might also be eligible to contribute to the population of 3-star validated super-centenarians. Some of these populations, like Japan, have super-centenarians on the full list but not on the 3-star validated list (this may change as the validation process moves forward). Since such cases are not included in the 3-star list of super-centenarians, it does not seem correct to include the population of these countries in this part of the analysis. At most, however, we might include a fraction of such populations. Thus, as an alternative scenario for the 3-star super-centenarians, we have added one-half of the population of the remaining countries in the Kannisto–Thatcher database, yielding a total of 376.7 million persons from whom the population of 3-star validated super-centenarians might be drawn, or about 42 times as large as the population of Sweden.

However, if the population were enlarged in this way, it would probably have a higher level of old-age mortality. Thus, we set the relative level of mortality at age 84 to 1.06, rather than 1.03. To match the data, we must increase r to around 1.15. In short, if the population that yielded the pool of super-centenarians were larger, then the mortality levels of this population would have to be higher in order to have yielded the observed data. However, we find the previous scenario more plausible and believe that the 3-star validated super-centenarians are drawn from a population that is only about 18 times as large as the Swedish population, with slightly higher mortality at ages 80–89 but substantially lower mortality above age 100 and especially above age 110.

Percentiles of the trend from the first (preferred) scenario for 3-star validated super-centenarians are shown in Figure 5. Remembering that the

FIGURE 5 3-star validated super-centenarians with predicted quantiles, birth cohorts of 1756–1884



NOTE: Predicted quantiles are based on the "3-star validated A" scenario of Table 4, which is our preferred model of the maximum age at death in the population from which such cases might arise.

data are truncated at age 110, the fit appears quite plausible. It is notable, however, that the case of Jeanne Calment (who died at age 122.44 in 1997) is truly exceptional. According to these calculations, the chance that the oldest member of her cohort (born in 1875) should live to this age or higher is a minuscule 2.19×10^{-4}. Thus, Jeanne Calment's exceptional longevity, relative to members of her own cohort, is the sort of event that should occur only about once every 4,500 years.[10]

Speculations about the world trend in maximum life span

Finally, we have tried to specify a scenario to depict the world trend in maximum observed life span. Our preferred scenario includes the population of all countries with at least one validated super-centenarian (whether in the 3-star or in a lesser category of validation). Presumably, this list includes most areas with mortality conditions and record keeping that are capable of producing validated world record longevity. As seen in Table 1, this group of countries has a population of 760.4 million people, or about 85 times the size of Sweden. We chose the same mortality parameters used in our preferred scenario for the 3-star validated super-centenarians. As shown in Table 4, this scenario leads to our speculation that the world trend in maximum life span may lie about 6–7 years higher than the Swedish trend.

Alternatively, we might include a larger population in this calculation, as shown in the second and third scenarios for the world in Table 4. In this case, we include all countries in either the super-centenarian or Kannisto–Thatcher databases, for a total population of 946.3 million, or about 106 times the Swedish population. However, it seems likely that such a larger population would have relatively higher levels of mortality at older ages. As seen in Table 4, if the mortality level is increased along with the population size, the median level of the maximum age at death does not increase and may even decrease. In short, for populations that are already so large, fluctuations in size play a relatively small role, whereas changes in mortality patterns are quite important.

As illustrated in Table 4, most of the difference in maximum life span between Sweden and the world is attributable to the difference in population size. Beneath the World I scenario, two additional rows show the median increase in maximum predicted life span if we vary population size or mortality levels alone. The 85-fold increase in population size, with no change in mortality, still yields more than a five-year increase in the maximum age at death for the most recent 40 cohorts. On the other hand, the difference in old-age mortality implied by the World I scenario (slightly higher among people in their 80s, but slightly lower among the very old) has little effect on its own.

The dominance of the population size effect is important for our final conclusion. We know that the world trend should lie about 5.3 years above the Swedish trend simply because of the larger size of the "eligible" population. However, we also know that a population 1,000 times as large as the Swedish population—thus larger than the current world population—would yield only about an 8-year increase in maximum life span. Absent a large difference in mortality between Sweden and the "eligible" population of the world (as defined earlier), the real difference should lie somewhere between these two extremes. Although mortality in Sweden is quite low by international standards across most of the age range, the evidence reviewed here suggests that Swedish mortality at very high ages (above age 100) is somewhat higher than in other industrialized countries. This mortality differential, when added to the population size effect, quite plausibly drives up the difference in maximum life span by about 6–7 years.

Conclusion

We have shown that it is possible to reconcile the Swedish trend in maximum observed life span with validated cases of super-centenarians in other populations, although the case of Jeanne Calment remains exceptional. The world trend in maximum life span is higher than in Sweden because of two factors: a larger population from which the world records are drawn, and atypically high mortality rates among Swedish centenarians. It is reasonable to speculate that the world trend in maximum life span may lie above the Swedish trend by about 6–7 years. If so, the maximum age at death in a calendar year would have been around 108 years in the 1860s (with most cases between 106 and 111) and about 115 or 116 years in the 1990s (with most cases between 114 and 118).

Notes

Several people gave helpful comments and suggestions about how to improve this chapter, including Kenneth Wachter, David Steinsaltz, and other participants in the Bay Area Colloquium on Population of September 2001. The authors also thank the many individuals who have contributed to this project by researching and validating the cases of extreme longevity discussed here, in particular Robert Bourbeau, Graziella Caselli, Dany Chambre, Bertrand Desjardins, Viviana Egidi, Louis Epstein, Bernard Jeune, France Meslé, Michel Poulain, Yasuhiko Saito, Axel Skytthe, Roger Thatcher, Jacques Vallin, and Robert Young.

1 Although England and Wales are part of the United Kingdom, they have historically maintained a separate statistical system. For this reason, it is convenient to treat them here as a separate country. Likewise, Scotland is presented separately in Table 1.

2 The model used here is correct whether the population is homogeneous or heterogeneous with respect to mortality risks. In a homogeneous cohort derived from N births, the probability that the maximum age at death is less than some age x is simply $F(x)^N$, where $F(x) = 1 - S(x)$ is the cumulative density function for age at death. This calculation, like all subsequent ones, assumes that the life spans of individuals are independent of one another (given an overall level of mortality).

On the other hand, assume that the above cohort is composed of k homogeneous groups of individuals, each having its own cumulative density and survival functions: $F_i(x) = 1 - S_i(x)$, for $i = 1,\ldots, k$. Let p_i be the probability that an individual chosen at random from the full population (of N births) belongs to group i, where

$$\sum_{i=1}^{k} p_i = 1.$$

Let X be a random variable denoting the survival time of a randomly chosen individual. It follows immediately that

$$F(x) = \Pr(X \le x) = \sum_{i=1}^{k} p_i F_i(x)$$

and

$$S(x) = \Pr(X > x) = \sum_{i=1}^{k} p_i S_i(x)$$

are the cumulative density and survival functions for the full, heterogeneous population. Observed death rates for the population as a whole can be used to derive estimates of $F(x)$ and $S(x)$.

In this model of a heterogeneous population, the sizes of the subgroups are not fixed in advance, but they must still sum to N. If Y_i is a random variable representing the size of group i, then the collection of group sizes, $(Y_1,\ldots, Y_k)$, has a multinomial distribution with parameters N and p_i, for $i = 1,\ldots, k$. For the full population, the probability that the maximum age at death is less than some age x is

$$\sum_{y_1, y_2, \ldots, y_k} [F_1(x)]^{y_1}[F_2(x)]^{y_2}\ldots[F_k(x)]^{y_k} \cdot \Pr(Y_1 = y_1, Y_2 = y_2, \ldots, Y_k = y_k),$$

where the summation extends over all possible combinations of group sizes. Substituting the formula for a multinomial probability in this equation, it is not difficult to show that the above expression equals $F(x)^N$.

Therefore, for this particular model of heterogeneity (which is itself quite flexible), it makes no difference whether we (i) compute probabilities associated with the maximum of the k maxima observed for homogeneous subgroups of the population (using death rates and survival curves computed separately for each group) or we (ii) compute such probabilities for the population as a whole, paying no attention to individual differences in survival. Using either method, the answer is identical.

3 Death rates were smoothed across cohorts to minimize the effects of random fluctuations.

4 The logistic model can be written

$$\mu(x,t) = \frac{r \cdot B(t)e^{\theta(t)x}}{1 + B(t)e^{\theta(t)x}},$$

where $\mu(x,t)$ is the force of mortality at exact age x for cohort t. The parameter, $B(t)$, defines the overall level of mortality relative to r, which is the upper limit to mortality at the highest ages. The increase of mortality with age is controlled by the parameter, $\theta(t)$. Some authors define the logistic model to be the above formula plus an additive constant, usually denoted as c (Thatcher et al. 1998). The additional term (also known as Makeham's constant) is useful when modeling mortality across the entire adult age range, but it is not necessary for studies of old-age mortality alone.

The logistic model has been recommended, in one form or another, for use in modeling old-age mortality because it is consistent with patterns observed in populations known to have high-quality data. Although there is no complete theoretical explanation for why the logistic curve fits old-age mortality patterns so well, it is possible that the process of physiological aging slows down considerably at very advanced ages (see Robine, this volume). An alternative explanation is that the logistic mortality curve is a result of population heterogeneity with respect to mortality risks. Specifically, if mortality for individuals follows a Gompertz pattern—but with proportional differences between individual mortality curves, where the constant of proportionality is distributed across the population according to a gamma distribution—then mortality for the population as a whole will follow a logistic pattern (see Thatcher et al. 1998).

Aside from these considerations, our choice to use only the logistic curve for modeling mortality at advanced ages in this study is based on a fairly simple observation: among the models of old-age mortality most commonly discussed in the literature, the logistic is the only one that both fits observed data well and does not yield impossible predictions about the level of mortality at very high ages. For example, the Gompertz model implies an overly rapid increase in old-age mortality and is thus inconsistent with known data. On the other hand, the quadratic model fits observed experience from age 80 to 110 quite well, but it yields decreasing and eventually negative death rates at higher ages.

5 Referring to Table 4 of Appendix D in Thatcher et al. (1998), it is possible to compute the upper asymptote of the mortality curve according to the formula

$$r = c + \frac{b}{\sigma^2},$$

where c, b, and σ^2 are parameters in the authors' version of the logistic model (which includes Makeham's constant, c, as mentioned in Note 4). After correcting an obvious typographical error in the value of one instance of σ, the range of r is consistently below 1 for all of the female data sets. Although it is higher than 1 for males in three out of four cases, it is the female mortality pattern that matters more when discussing extreme longevity for a population as a whole.

6 Kannisto proposed that observed death rates at older ages, M_x, adhere closely to a linear model after performing a logit transformation:

$$\ln\left(\frac{M_x}{1 - M_x}\right) = a + bx + \varepsilon_x,$$

where ε_x is an error term.

7 The raw material of the Kannisto–Thatcher Database on Old-Age Mortality consists of death counts at ages 80 and older classified by sex, age, year of death, and year of birth. The database also provides age-specific annual population estimates (on 1 January) over the full range of the data series for each country. These population estimates were derived from raw death counts using the methods of extinct cohorts (Vincent 1951) and survivor ratios (Thatcher et al. 2001). Such information was used in the current study to compute death rates over age and time for individual countries or groups of countries.

8 For purposes of discussion and comparison, it is convenient to express the level of mortality in the logistic model in terms of mortality around ages 80–89. Therefore, in Table 4 and elsewhere, we use $\mu(84,t)$, the force of mortality at exact age 84 for cohort t, instead of the parameter, $B(t)$, that appears in the formula for the logistic curve in Note 4. It is easy to verify that

$$B(t) = e^{-84\theta(t)} \frac{\mu(84,t)}{r - \mu(84,t)}.$$

9 For the scenarios involving the "complete and validated" countries, or larger groups that include this set of countries, we have counted Belgium in calculations of population size, even though we lack comparable mortality data for this country above age 100 and have omitted Belgium from Tables 2 and 3. In the super-centenarian database, Belgian cases have never attained world-record status even for a single calendar year. Given this fact plus its small size, the choice to include or exclude this country in these rough calculations of relative population size has no relevance to our conclusions.

10 If the trend in maximum life span continues to increase, it is possible and even likely that Jeanne Calment's record longevity will be surpassed within the next few decades. It is important to emphasize that the probability given here estimates the chance of observing such an extreme value for an individual born in 1875, not in any other year.

References

Aarssen, Karin and Laurens de Haan. 1994. "On the maximal life span of humans," *Mathematical Population Studies* 4(4): 259–281.

Brillinger, David R. 1986. "The natural variability of vital rates and associated statistics," *Biometrics* 42(4): 693–734.

Gumbel, Emil J. 1937. *La durée extrême de la vie humaine*. Paris: Hermann.

———. 1958. *Statistics of Extremes*. New York: Columbia University Press.

Jeune, Bernard and James W. Vaupel (eds.). 1999. *Validation of Exceptional Longevity*. Odense, Denmark: Odense University Press.

Kannisto, Vaïnö. 1994. *Development of Oldest-Old Mortality, 1950–1990: Evidence from 28 Developed Countries*. Odense, Denmark: Odense University Press.

Robine, Jean-Marie and James W. Vaupel. 2001. "Supercentenarians: Slower ageing individuals or senile elderly?" *Experimental Gerontology* 36(4-6): 915–930.

Thatcher, A. Roger, Vaïnö Kannisto, and Kirill Andreev. 2001. "The survivor ratio method for estimating numbers at high ages," *Demographic Research* [Online], Vol. 6, available at http://www.demographic-research.org/volumes/vol6/1

Thatcher, A. Roger, Vaïnö Kannisto, and James W. Vaupel. 1998. *The Force of Mortality at Ages 80 to 120*, Odense, Denmark: Odense University Press.

Vincent, Paul. 1951. "La mortalité des vieillards," *Population* 6(2): 181–204.

Wilmoth, John R., Leo J. Deegan, Hans Lundström, and Shiro Horiuchi. 2000. "Increase of maximum life-span in Sweden, 1861–1999," *Science* 289: 2366–2368.

Wilmoth, John R. and Hans Lundström. 1996. "Extreme longevity in five countries: Presentation of trends with special attention to issues of data quality," *European Journal of Population* 12(1): 63–93.