IDL, the International Database on Longevity

Jean-Marie Robine*, Jutta Gampe**, Amandine Cournil* and James Vaupel**

**Health and Demography*, INSERM, Montpellier, France, <robine@valdorel.fnclcc.fr>; **Max Planck, Institute for Demographic Research, Rostock, Germany

Abstract: Since the first symposium, "Living to 100 and Beyond: Survival at Advanced Ages" held in 2002, a collaborative effort has been made to assemble an international database on longevity, gathering validated longevity records for people having reached at least their 110th birthday. More than 15 countries, including the United States of America, Canada and Japan, along with European countries, have been participating in this Supercentenarian project. Collaboration with national statistical offices or health departments has allowed the investigators to obtain complete lists of alleged supercentenarians in most countries. Different validation processes are then undertaken by the participating teams. By March 2004, more than 500 validated records had been gathered. This paper first evaluates the quality of data according to several criteria, such as the country of residence or the validation process undertaken, and then provides an estimation of the mortality trajectory up to age 114.

Introduction

Although the Gompertz exponential model has for a long time been the prevailing method used to explain the increase in death rates in human populations, evidence indicates that the increase in the risks of death tends to decelerate at older ages (something that the Gompertz model does not account for). Different patterns of mortality deceleration have been proposed, ranging from a slower increase for the oldest of the old to a rapid decrease of mortality rates above a certain age (Vaupel *et al*, 1998; Thatcher, 1999). However, sufficient and reliable demographic data were not available for us to be able to discriminate according to these different hypotheses. This context has motivated an international collaborative effort to gather demographic data for the oldest ages in order to create the International Database on Longevity, or IDL (Robine and Vaupel, 2002).

The International Database on Longevity

The aim of this database is to compile complete and validated lists of supercentenarians (defined as those 110 years old and over) as reliable data to estimate mortality trajectories. These data are provided by statisticians and demographers involved in IDL as contributors (www.supercentenarians.org). The participating countries include 13 European countries (Austria, Belgium, Denmark, Finland, France, Germany, Italy, the Netherlands, Norway, Spain, Sweden, Switzerland and United Kingdom), as well as Japan, Australia, and North America (the United States and Canada). The database currently contains validated cases from 14 countries, including 6 countries where complete (or age ascertainment bias free) lists are available.

The term supercentenarians refers to a living or dead person who has reached, at minimum, the age of 110 years. The IDL currently focuses on age 110 and over. However, future development of the database might also include younger ages, down to 106 for women and 105 for men[1]. This would ensure a complete matching with the Kannisto-Thatcher mortality database, which covers ages 80 to 108 (Kannisto, 1996; Thatcher *et al*, 1998).

Reliable estimation of mortality trajectories for such extreme ages should avoid two kinds of bias: incompleteness and errors in the reported ages. A lack of completeness may be a source of bias if identification of the cases is not age-independent. Media coverage, for example, is usually best for the oldest cases. Thus press reports do not generally constitute an appropriate source of data for mortality trajectories estimation. IDL should then be supplied with complete lists as much as possible, where complete means that all supercentenarians have been reported and identified in a given country and for a given time-period. Incompleteness may be overcome if the list results from an age bias-free method of identification. Age bias-free lists are suitable to estimate rates of mortality and will generally be associated with complete data.

Validation

Studying extreme ages necessarily involves small numbers and, therefore, a single age-error, especially at the highest ages, may well have strong implications for the pattern of the

[1] Already available for France.

mortality trajectory. It is thus of great importance to check the reliability and accuracy of each reported age from such a dataset. The method employed to check age is referred to as the validation procedure and has been a major issue all throughout the development of the database. The validated list only includes supercentenarians for which the reported age has been ascertained and checked (Cournil *et al*, 2005). The validation procedures differ among the various countries involved according to the availability of resources used to check the age. In order to take into account the different situations, IDL contains a validation variable which classifies the cases according to the information that was used to consider the case as validated. Three main situations are encountered. First, the case is not considered to be validated as the reported age has not been sufficiently verified, or ascertained (level of validation = 0). Another situation occurs when the case is considered validated and the main type of documents which were used to perform the validation are recorded in the database (level of validation = 2). There is a third in-between situation where the case is considered to be validated, but no detailed information is available in the database (level of validation = 1). The IDL relies on a reference person to consider the case validated. The reference person or team, also called the country informant(s), can be in charge directly or indirectly of the validation of the case and of its report or transmission to the IDL. In some countries, the list of cases transmitted to the IDL comes from an institutional source, such as a statistics institute or an administration. The source of information is included in the database. For cases validated by documents, additional variables describe the type of source which has been used to ascertain the date of birth (i.e., birth record or census documents) and the date of death. Another field denotes whether the date of birth is attested to by a photocopy, photo or scan of an original document.

Completeness

The IDL currently gathers validated and complete (or age-ascertainment bias-free) lists of supercentenarians from 6 countries, including Denmark, France, Italy, Japan, the United Kingdom (England and Wales), and the United States, as well as from one province of Canada (Quebec). For all these countries (and this one province), lists are provided through collaboration with a specific statistics institute or an administration: Civil Registration System (CRS) for Denmark, Institut National de la Statistique et des Etudes Economiques (INSEE) for France, Istituto nationale di statistica (ISTAT) for Italy, Ministry of Health and Welfare (MHW) for Japan, Office for National Statistics (ONS) for England and Wales, Social Security Administration (SSA) for USA and Institut Statistique du Québec (ISQ) for the province of Quebec (Canada). These institutes provide lists of deaths, except in Japan, where the MHW provides the list of surviving centenarians on the 1st September of each year. The lists from France, Italy, Japan, Denmark, England and Wales, and Quebec are assumed to be complete as they rely on an exhaustive report of cases. Their completeness concerns a given time-period for which data were available. The list transmitted to the IDL from the SSA for the USA is not complete because it only includes cases which have been validated. However, this list is assumed to be age ascertainment bias-free (Kestenbaum and Fergusson, 2002; Rosenwaike and Stone, 2003). The supercentenarian cases of these countries have all been submitted to a validation procedure[2]. Four of the countries, Canada, France, Italy and the United States, have provided validation information for each single case. This information mentions the main documents (vital records, censuses, etc.) that were used to check for the dates of birth and death, and they are available in the database. Paper files can also be consulted for supercentenarians in various locations (mainly from Montpellier and Duke Universities).

Validation is based on cross-checking birth and death records (or their proxy such as an early census document) for each reported case. The validation procedure for the whole list has allowed the identification of false supercentenarians in the initial lists provided by the French, Italian and Quebec provincial statistical bureaus. The knowledge of the number and type of errors gives information about the quality and accuracy of the national registration systems. It may argue for or against the need for a validation procedure for each single case. Eight (8), five (5) and seven (7) errors were found in France, Italy and Quebec respectively, corresponding to the error percentages of 14%, 19% and 46%. These high values argue in favor of individual verification of the reported ages. The lists also need to be evaluated for their completeness, but this is very difficult to do. One method is to compare lists of supercentenarians provided by different sources. In France, the comparison of the list provided by INSEE with other lists obtained by different means identified five additional cases that were not mentioned in the INSEE list. These cases mainly come from the first years of the observation period. It seems that the quality of the list in terms of completeness has improved with time. Basically, to assume completeness of the lists, we have to trust the statistics institutes in charge of the registration of vital events (Cournil *et al*, 2005).

[2] Japan is an exception, as only 29 cases out of the 96 have been submitted to an individual validation procedure. These 29 cases correspond to dates of death between 1995 and 2000.

Five other lists, from Belgium, the Netherlands and three Nordic countries (Finland, Norway and Sweden) are assumed to be complete. However, for these countries, information that would be needed to ascertain the reliability of the data, both on the validation procedure and completeness, is lacking. Cases for Belgium and the Netherlands seem to be known from an extensive press review. Data from the three Nordic countries come from a linkage between statistic input and press reports. A common feature of these countries is that they are submitted to very strict legislation concerning access to personalized information, which renders very difficult the dissemination of nominative lists for supercentenarians and also the access to official civil documents to check and validate known cases from press reports (Cournil *et al*, 2005).

Gathering the four datasets

Taking into account all these characteristics, we have divided the data gathered in the IDL into four datasets according to their level of validation and degree of completeness.

The level of validity

As of March 2004, the international database on longevity (IDL) contained 732 records corresponding to 78 males and 654 females, alive or dead, alleged to have reached the age of at least 110 years. Most of them, 677 records (92%), correspond to deceased individuals (72 males and 605 females).

For 401 (or 59%) of the deceased individuals, copies of genuine documents corroborating their dates and places of birth and death (birth record or proxy such as a census sheet or death record) are available. For 147 individuals (22%), a special check has been made by the statistical office providing the data or by the IDL informant, but copies of the genuine documents are not yet available. For 129 individuals (19%), no special validation has been undertaken by the statistical offices or by the IDL informants (see Table 1). Some of these unvalidated cases came from previous lists of supercentenarians and even the country of death may be unknown.[3]

Table 1 about here

[3] See the Los Angeles Gerontology Research Group for such lists (Coles, 2004).

The country of death

Indeed, the country of death is known for only 660 individuals (97%), involving 21 countries in total. Of this number, a mere four countries gathered fully 84% of the data: the USA with 352 cases (53%), Japan with 75 cases (11%), France with 66 cases (10%) and England and Wales with 61 cases (9%). Conversely, seven countries provided only one case each (see Table 2).

Table 2 about here

The level of completeness

The lists of deceased supercentenarians are expected to be complete for 11 countries as well as in the Canadian province of Quebec. For six countries, the United States, France, England and Wales, Japan, Italy, Denmark, and the province of Quebec, the lists came from institutional sources such as the Social Security (USA), Statistical Bureaus (France, England and Wales) or the Ministry of Health (Japan). For an additional five European countries, Belgium, Finland, the Netherlands, Norway and Sweden, even if the lists are not provided by an institution, they are expected to be complete thanks to the existence of population registers and wide coverage of the supercentenarians by local media.

The four datasets

Using the country of death, the source of data and the level of validation, we assembled four datasets (see Table 3). The first dataset comprises the lists that come from institutional sources.

For the United States, the data come from the Social Security Administration (SSA) with a level of validation of 2 for 326 cases, deceased between 1980 and 2003 (Kestenbaum and Fergusson, 2005).

For France, the data come from the French statistical bureau (INSEE) and have been validated by the Health and Demography Research Group at the University of Montpellier, which leads to a level of validation of 2 for 49 cases, deceased between 1987 and 2002 (Meslé *et al*, 2005).

For England and Wales, the data come from the statistical bureau (ONS) which makes a special check of the data before releasing them annually to the IDL, leading to a level of validation of 1 for 53 cases, deceased between 1968 and 2002.

For Japan, the data come from the Ministry of Health (MHW) which annually releases the list of centenarians living in Japan. A special check made by Nihon University Population Research Institute led to a level of validation of 1 for 29 cases, deceased between 1996 and 2001 (Saito, 2005).

For Italy, the data come from the Italian statistical bureau (ISTAT). A validation of the data by the Department of Demography of the University of Rome, La Sapienza, is in progress, leaving the level of validation unknown for 21 cases, deceased between 1973 and 2000 (Barbi *et al*, 2005).

For Denmark, the data come from the National Civil Registration System (CRS) with a special check of the data made by the University of Southern Denmark, leading to a level of validation of 1 for 2 cases, deceased between 1996 and 2000 (Skytthe *et al,* 2005).

For the province of Quebec, Canada, the data come from the statistical bureau of Quebec (ISQ) and have been validated by the Department of Demography at the University of Montreal, leading to a level of validation of 2 for 6 cases, deceased between 1983 and 2001 (Desjardins and Bourbeau, 2005).

In total, this first dataset comprises 486 records (42 males and 444 females), coming from 7 countries with a level of validation of 2 for 381 cases (USA, France and Quebec), 1 for 84 cases (England and Wales, Japan and Denmark) and unknown (in progress) for 21 cases (Italy).

Table 3 about here

The second dataset comprises the lists coming from five European countries where the lists are expected to be complete, even in the absence of institutional sources. Only the cases which have been the subject of a special check by IDL informants have been kept, leading to a level of validation of 1 for a total of 25 additional cases (4 for the males and 21 for the females), deceased between 1857[4] and 2001 (16 cases in the Netherlands, 5 cases in Belgium, 2 cases in Finland, 1 case in Norway and 1 case in Sweden) (Skytthe *et al*, 2005; Poulain, 2005).

The third dataset comprises 36 records (6 males and 30 females) from six countries with a validation level of 1 for 16 cases and 2 for 20 cases. Half of these cases come from countries where we have no access to institutional lists and where the number of validated or known cases are too low, such as Australia (7 cases, see McCormack, 2005), Canada (11 cases, see Desjardins and Bourbeau, 2005) and Ireland (1 historical case). Half of the cases come from countries where we have access to institutional lists. These may be missing or recent cases such as in the USA (1 case) or Italy (2 cases) or they may be missing or old cases that were recorded prior to the institutional list such as in France (14 cases, see Meslé *et al*, 2005).

The fourth dataset is made of 113 records (including 18 males and 95 females), with a level of validation of 0, from 16 countries, including 46 cases from Japan that need to be validated. A part of this dataset overlaps with the lists made by the Los Angeles Gerontology Research Group.[5]

---

[4] The Dutch list comprises three historical records, corresponding to two men born in 1745 and 1788 and one woman born in 1849, which are still debated by our Dutch and Belgium informants.

[5] The Los Angeles Gerontology Research Group (GRG) draws up nominative lists of known supercentenarians, mainly collected through the media (www.GRG.org). The GRG makes the names of the supercentenarians publicly available, in particular the names of the living supercentenarians. Contrary to the IDL, the GRG does not place great importance on the completeness of its lists. Their main lists are List A made by Louis Epstein (Chronological listing of all supercentenarians) and List AA made by Robert Young (Chronological listing of all supercentenarians with Nationality). The GRG list of living supercentenarians (List E) contained 61 cases as of March 11, 2005. In comparison to the IDL, the GRG mainly contains cases from countries not yet participating in the IDL and old cases prior to the periods covered by the IDL lists or, on the other hand, very recent cases (including the living centenarians) for which the IDL does not yet have the information from the statistical bureaus.

The supercentenarians

Emergence and rate of increase

The first supercentenarians, except for a few historical cases, appeared in the 1970s[6]. Since 1980, their number has increased at least exponentially until 1999, the last year for which we collected data from all the institutional sources (dataset 1 on Figure 1). Dataset 4 (which is not yet validated) shows a similar exponential increase until the most recent years, suggesting that this remarkable increase is still ongoing (Figure 2).

Figures 1 & 2 about here

Figure 3 details this increase by institutional source for the seven countries which provide official lists (dataset 1). Only the US list is large enough to display a clear increase over time (labelled SSA on Figure 3). Although the global increase is obvious, some countries such as England and Wales (labelled ONS on Figure 3) display no significant trend.

Figure 3 about here

Age structure

Figure 4 shows that the age structure of the supercentenarians varies according to the datasets. In datasets 1 and 2, which are thought to be complete, the individuals deceased at age 110 are more numerous (51% and 64%) than in dataset 3, which is known to be incomplete, or than in dataset 4 (33% and 45%, respectively). This confirms the hypothesis that when a list of supercentenarians is incomplete, the probability of missing cases decreases when the age of the supercentenarians increases. Indeed, the higher the age reached, the larger is the probability to be known. Conversely, datasets 3 and 4 comprise many more individuals alleged to be dead at age 114 or more (11% in both datasets), compared to datasets 1 and 2 (6% and 0%, respectively). The most obvious explanations are the better awareness of the oldest old for dataset 3, the absence of validation and potentially false cases for dataset 4, and the small size of dataset 2. Only the first dataset gathers the three necessary elements for

[6] In particular, Robert Young is currently documenting a few cases of early American supercentenarians, such as Delina Filkins (1815-1928) and Betsy Baker (1842-1955) (personal communication).

assessing the age structure of the supercentenarians, i.e., the completeness of the lists, their validation and a sufficient size.

Figure 4 about here

Sex-ratio

Again, the sex-ratio varies according to the datasets (Figure 5). In the first dataset, males represent only 9% of the record versus about 16-17% in the three other datasets. Similar explanations can be provided: a better awareness of the male supercentenarians for the incomplete dataset 3, an absence of validation and possible false cases for dataset 4 and the small size and small numbers for dataset 2. The first dataset probably provides the best estimation of the true sex-ratio.

Figure 5 about here

Maximum age at death

The emergence of the supercentenarians was accompanied by a significant increase in the Maximum Reported Age at Death (MRAD). Some have discussed a handful of historical cases of individuals alleged to be dead at age 110 or age 111 before 1960, but no cases of death at age 112 or over have been proven to exist before 1973. Since that time, however, deaths have been recorded at age 114 in 1985, age 117 in 1993 and age 122 in 1997 (Figure 6).

Figure 6 about here

The mortality trajectory at age 110 and over

As the records in the IDL contain information on deceased persons only (Japan excepted), supercentenarians still alive are not part of the dataset. To correctly estimate the number of persons at risk at different ages, the numbers of survivors beyond each age have to be estimated from death counts for those cohorts that are not yet extinct. This can be achieved by the Survivor Ratio Method which has proved to be the most reliable technique among several available options (Thatcher *et al*, 2002; Andreev *et al*, 1999). The Survivor Ratio Method extends the method of extinct generations (Vincent, 1951) to ‘almost extinct’ cohorts. In this approach specific ratios of survivors in a cohort beyond some age relative to the number of deaths in this cohort during the previous k years are assumed to be stable across neighboring cohorts. By starting from extinct cohorts the number of survivors can be estimated in more recent cohorts. If death rates are changing (falling), then survivor ratios are not constant from one cohort to the next. In this case the introduction of a correction factor can compensate for the changing survivor ratios (Thatcher *et al*, 2002).

Before applying the survivor ratio method to the more recent cohorts in the database, we checked whether mortality experience has changed over time for supercentenarians by analyzing data for those cohorts which are already extinct. Ages at death for individuals in the IDL are available up to the date, allowing us to apply methods for continuous survival data. In order not to mix patterns of different countries, we have initially restricted our analysis to the US contribution to the IDL. Because of its sample size, the list from the SSA (USA) allows statistically reliable conclusions. Analyses were conducted separately for males and females.

To determine which cohorts have a completed mortality experience we had to fix the age ω beyond which no member of a cohort is alive. A natural choice might be the maximum (validated) age we are aware of, which would be 122 year for females and 115 for males. On the other hand, for more recent cohorts and in the countries contributing to the complete and validated list in the IDL, we can be confident that we would be aware of any individual currently alive at an age greater than 115. For this reason we picked ω=116 as the maximum age used to define already extinct cohorts. With this setting, the birth cohorts of 1887 and earlier have completed their mortality experience according to the US sampling scheme.

To investigate the basic survival pattern and to check whether this pattern has changed for earlier versus later cohorts, in a first step we split these cohorts into an earlier group, 1870-1880, and a later group, 1881-1887. The split was made in such a way that the two groups are of comparable size. Figure 7 shows the two estimated survival probabilities (Kaplan-Meier estimates) for US females in these two groups as closely coinciding. A log-rank test supports the similitude of the two survival patterns (p>0.7).

Figure 7 about here

Another striking feature of these survival curves is the seemingly exponential decline in survival probabilities. To check for the appropriateness of an exponential distribution, we fitted the parameter by maximum likelihood to the whole sample of extinct cohorts ($\hat{\lambda} = 0.764$). Figure 8 shows the nonparametrically estimated survival probabilities as well as the integrated hazard (confidence limits for $(1-\alpha) = 0.95$). The estimated exponential distribution is completely contained within the confidence bounds. A two-sided Kolomogoroff-Smirnov test (D=0.0299, p=0.984) confirms the exceptionally close fit of the data to an exponential distribution. As a consequence, the force of mortality beyond age 110 is constant at the level $\mu = 0.764$, corresponding to an expected residual life span of 1.309 years. The resulting estimated annual probability of death is 0.534.

Figure 8 about here

The total number of male individuals in the extinct US cohorts is only 23 and thus does not allow for further division. The parallel analysis for males is documented in Figure 9. Again the data strongly support a constant hazard (KS-Test D= 0.0853, p= 0.996), which for males is estimated as $\mu = 0.608$, corresponding to an expected residual life time of 1.64 years and an annual probability of death of 0.456.

Figure 9 about here

After exclusively analyzing the US data, all individuals in extinct cohorts from data sets 1 and 2 (see above) were considered. The data involved in these sets are sufficiently complete and validated. Figures 10a and 10b give the corresponding results for these combined cases, which basically conform to the analyses of the US data. Both for male and female supercentenarians, the hypothesis of a constant hazard beyond age 110 is distinctly supported. Nevertheless, the agreement is slightly less pronounced in these datasets than in the US data alone (KS-Test for females: p=0.597; for males: p =0.668). We are currently in the process of investigating whether some countries/cohorts exert a particular influence on the overall fit.

Figures 10a & 10b about here

Using the survivor ratio method, the number of individuals who were exposed to risk but unobserved because they have not yet died is estimated for the most recent cohorts. The results of the final analysis are presented in Figures 11a and 11b. Again, constant hazards are strongly supported up to ages at death of about 114. After this, age data grow sparse and hence the confidence limits widen considerably. The estimated constant hazard is $\mu = 0.744$ for females and $\mu = 0.664$ for males, leading to a remaining life expectancy of 1.34 and 1.51 respectively. Annual probabilities of death for these hazards are 0.525 and 0.485 respectively.

Figure 11a & 11b about here

Discussion

The mortality trajectory of human aging is a major biomathematical question that has fascinated demographers and biologists for a very long time. Is there a law of mortality? And if so, what is it? It has been known for some time that mortality increases with age. But is there an age at which death the following year is a certainty? And, if yes, what it is? Research into laws of mortality has been marked by foundational work by English mathematicians, particularly Moivre (1725), on the increase of mortality with age; Gompertz (1825), on the exponential progression of mortality; and Perk (1932), on a so-called logistical trajectory. But this area of inquiry also owes a great deal to biologists who, since Buffon, have been pointing out that the mortality of old people is lower than expected, a phenomenon also recognized by Gompertz himself (Gompertz, 1825; Carnes *et al*, 1996; Olshansky and Carnes, 1997).

Thus Buffon, as early as 1749, noted that the yearly probability of surviving with age "*decreases less and less, [and] becomes stationary and fixed*" when the lifespan is complete; i.e., at about eighty years, according to him. Gompertz made a similar observation in 1825 from Milne's table of Carlisle, stating that, "*Such a law of mortality would indeed make it appear that there was no positive limit to a person's age.*" In 1939, Greenwood and Irwin again speculated on the "*possibility that with advancing age the rate of mortality asymptotes to a finite value*." They added, "*One cannot without absurdity believe that, other things equal, a man of 100 is not more likely to die within a year that a man of 90. But other things are not equal…" (Greenwood and Irwin, 1939).* Eventually, in 1964, Comfort noted, "*Extreme records in man, occurring in excess of statistical probability, are chiefly of interest in suggesting that after a certain age the rate of increase in the force of mortality is not maintained, either by reason of selection or from other causes.*" (Comfort, 1964). The dates of publication of the previous works, 1749, 1825, 1939, and 1964, respectively, testify that the hunch that the probability of dying within a year can reach a plateau is quite old, but that it is still difficult for most observers to believe. Since 1932, the logistic model developed by Perk has allowed a fitting of such trajectories. But the exponential trajectories, following the model of Gompertz (1825) or the modified one by Makeham (1860), have remained the preferred models until quite recently and are still in use today in some statistical bureaus and actuarial institutes. However, it is now well recognized that mortality slows down beyond the age of 85 and that mortality rates observed at 95 or 100 years of age are lower than expected with a Gompertz-like trajectory (Kannisto, 1996; Horiuchi and Wilmoth, 1998; Thatcher *et al*, 1998; Vanfleteren *et al* 1998; Vaupel *et al*, 1998 ; Kannisto, 1999; Manton, 1999; Thatcher, 1999; Lynch and Brown, 2001; Weitz and Fraser, 2001; Kestenbaum and Fergusson, 2002; Barbi *et al*, 2003). Still it is unclear whether a plateau of mortality is reached and, if so, at which age and at which level.

From 1951 onwards, the extinct generation method suggested by Vincent (1951) and its modern extensions (Andreev *et al*, 1999, Meslé and Vallin, 2000; Thatcher *et al*, 2002) were used to calculate the mortality rate for the oldest-old cases accurately. The subsequent creation of international databases covering several countries has increased the number of observations (Vincent, 1951; Dépoid, 1973) on a limited basis until the creation of the Kannisto-Thatcher Database (KTD), which includes mortality data from 28 developed countries (Kannisto, 1994). This database, now associated with the Human Mortality Database (www.mortality.org or www.humanmortality.de), is the foundation for all modern work on mortality trajectories which suggest either logistic trajectories (Kannisto, 1996; Thatcher *et al*, 1998; Thatcher, 1999) or quadratic trajectories (Vaupel *et al*, 1998). The International Database on Longevity (IDL) was set up predominantly to be able to distinguish between these two trajectories. At minimum, its objective is to accurately calculate mortality rates between the ages of 105 and 115 years old (Robine and Vaupel, 2002).

Since the work of Gompertz (1825), the entire scientific community (excepting the few scholars quoted above) has believed that mortality increases exponentially with age, with mortality rates doubling for humans approximately every eight years. In 1950, Vincent was the first to be able to estimate the age at which the probability of dying in the same year was close to 1: his estimate was 110 years (Vincent, 1950). Twenty years later, based on the same hypotheses and methods, Dépoid (1973) found a limit at the age of 117 years old. This collapse in mortality is considerable, but at the time hardly anybody was interested in this question and even subsequently doubts have never been raised over Vincent's conclusions. The duration of human life is limited to 110 years – this is a characteristic of the species (Buffon, 1749; Cutler, 1985; Walford, 1985; Hayflick, 1996) which is controlled by clock genes (Fries, 1980; Hayflick, 1981). The Kannisto and Thatcher database had to be created before this biological dogma was questioned. Mortality follows a logistic trajectory with age, tending towards a mortality ceiling without ever reaching it (Thatcher *et al*, 1998), or a quadratic trajectory, where mortality diminishes with age after having passed a maximum at around 110 years old (Vaupel *et al*, 1998). The mortality rates between 110 and 114 years of age, calculated accurately thanks to the IDL and presented in this paper, show that mortality does not increase, or only very slightly, after 110 years of age. These observations definitively reject the exponential trajectories (Gompertz or Makenham), but still do not make the distinction between the logistic and quadratic trajectories, with the latter only diverging significantly after 115 years of age (Robine and Vaupel, 2002). These results nevertheless show that human longevity has no limits in terms of age and there are, accordingly, no biologically-controlled limits (such as clock genes or other mechanisms linked to natural selection). If there really is a limit, it is for mortality levels that do not reach an annual mortality rate of 60%, even in people of 110 years of age or more, even for the frailest elderly people, under mortality conditions currently found in the most developed countries.

These results require some explanation, for they fuel the fear that old people will always be both increasingly numerous and increasingly older. What determines the mortality ceiling? Could it drop significantly in the future? Several explanations are put forward to explain the slowing down in mortality of the oldest-olds, like genetic heterogeneity of populations (Vaupel *et al*, 1979), the redundancy of biological maintenance, repair and defence mechanisms in the human body (Gavrilov and Gavrilova, 2001 and 2003), the existence of a very special configuration of generations at the heart of the demographic transition (Bonneux *et al*, 1998), considerable progress in the economic, physical, medical and social environments of the oldest people (Robine, 2001 and 2003), and the plasticity of mortality to such changes (Vaupel *et al*, 2003). In 1939, Greenwood and Irvin already noted that, "*Even the juvenile of 60, if ordinarily intelligent, eschews the violent exercises of the child of 40. Centenarians rarely appear in public. A statistical rate of mortality might show no increase with age, if the demands made on the vital forces diminished pari passu with the decay of vigour*" (Greenwood and Irwin, 1939). This question was entirely theoretical in 1939, but in 2005, with the accumulation of nonagenarians and centenarians either housebound or in retirement homes, the determining factors of mortality trajectories with age should be investigated urgently in order to determine the level to which mortality ceilings can be lowered. Remember, for example, that in the 25 years between 1975 and 2000, the annual probability of death for women centenarians in Japan decreased from 50 to 35% (Robine, 2003).

Already in 1939, Greenwood and Irwin proposed a limiting value of 0.5 for qx "*using data from 1920-1922*." This suggests that there has been no change in the mortality ceiling for the last 80 years but we do not know which kind of data they used or the quality of these data. By contrast, the IDL gathers complete and validated lists of supercentenarians from the main low mortality countries in Europe and North America as well as from Japan and Australia, with the collaboration of the national statistical bureaus. In March 2004, the IDL already contained 677 records corresponding to deceased individuals, with most of the deaths having occurred after 1995. Furthermore, the numbers have rapidly increased in the participating countries and additional countries, such as Austria, Germany (Maier and Scholz, 2005), Spain and Switzerland, should soon be providing their own lists to the IDL.

In this study, we have shown that constant hazards are strongly supported by the IDL from the age of 110 years to the age of 114 years, with annual probabilities of death close to 0.5. We expect in the next few years to be able to extend our calculations to the ages of 115 to 116 years, ages at which the logistic and quadratic trajectories start to significantly diverge.

The study of mortality trajectories is one domain of bio-demography in which work on animal species complements work on humans (Carey *et al*, 1992; Curtsinger *et al*, 1992, Carey 1997 and 2003). Death decelerates with age for insects and worms as well as for humans (Vaupel *et al*, 1998). Mortality deceleration can be an artefact of compositional change in a heterogeneous population (Vaupel *et al*, 1979; Brook *et al*, 1994). Therefore, genetic heterogeneity could explain similar mortality trajectories in human populations in which mortality stops increasing exponentially beyond 85 years of age. But data from *Drosophila* flies demonstrate that a levelling off of death can occur even when heterogeneity is minimized through the rearing of genetically homogeneous cohorts under similar conditions (Vaupel *et al*, 1998) and that Gompertz's law does not hold for either the isogenic or heterogeneous nematode populations (Vaupel *et al*, 1994). Post-reproductive life-spans might be compared with a '*post-warranty survival of equipment*' (Vaupel, 2003). Although living organisms are much more complex than manufactured products, the trajectory of the mortality of automobiles (Vaupel et al. 1998) suggests that deceleration is a general property of complex systems. This proposal is similar to the reliability theory of aging and longevity proposed by Gavrilov and Gavrilova (2001 and 2003). However, outside the laboratories, all populations are heterogeneous and "*the frail tend to suffer high mortality, leaving a subset of survivors,*" which creates a fundamental problem in the analysis of mortality trajectories. The central question remains, "*how important are an individual's survival attributes as opposed to current conditions in determining the chance of death?*" (Vaupel *et al*, 1998).

Acknowledgements

The authors want to thank Leonid Gavrilov for the discussion of a previous version of this paper during the meeting of the "Society of Actuaries, Living to 100 and Beyond" held in Orlando in January 2005, and for introducing us to the paper of Greenwood and Irwin published in 1939.

The authors also want to thank the IDL informants, E. Barbi, S. Bruzzone, G. Caselli, G. Desplanques, B.R. Fergusson, A. Hervonen, D. Hoelseth, B. Jeune, B.M. Kestenbaum, H Maier, J. McCormack, F. Meslé, M. Poulain, Y. Saito, R. Scholz, A. Skytthe, A.R. Thatcher, F. Tillo, J. Vallin, and R. Young, as well as the statistics institutes and administrations providing the data: the Civil Registration System (CRS) for Denmark, Institut National de la Statistique et des Etudes Economiques (INSEE) for France, Istituto nationale di statistica (ISTAT) for Italy, the Ministry of Health and Welfare (MHW) for Japan, the Office for National Statistics (ONS) for England and Wales, the Social Security Administration (SSA) for USA and Institut Statistique du Québec (ISQ) for the province of Quebec (Canada). Eventually, the authors want to thank Jenae Tharaldson for her assistance in the editing.

References

Andreev, K. (1999) *Demographic Surfaces: Estimation, Assessment and Presentation, with application to Danish Mortality, 1835-1995*. Ph.D. thesis, University of Southern Denmark.

Barbi E, Gaselli G, & Vallin J. (2003) Trajectories of extreme survival in heterogeneous populations, *Population* E58 (1): 43-66.

Barbi E, Bruzzone S & Caselli G. (2005) *Updated collection of Italian deaths over 110 by region, sex and birth cohort.* Background paper for the IDL Monograph (To be published).

Bonneux L, Barendregt JJ, & Van der Maas PJ. (1998) The expiry date of man: a synthesis of evolutionary biology and public health. *J Epidemiol Community Health* 52:619-23.

Brooks A, Lithgow GJ & Johnson TE. (1994) Mortality rates in a genetically heterogeneous population of "Caenorhabditis elegans". *Science* 263:668-671.

Buffon. (1749) *Oeuvres Complètes*. Paris: P Duménil Editeur (Edition de 1836).

Carey JR, Liedo P, Orozco D, & Vaupel JW. (1992) Slowing of mortality rates at older ages in large medfly cohorts. *Science* 258:457-461.

Carey JR. (1997) What demographers can learn from fruit fly actuarial models and biology. *Demography* 34(1):17-30.

Carey JR. (2003) *Longevity, the biology and demography of life span*. Princeton: University Press; 2003.

Carnes B, Olshanski SJ, & Grahn D. (1996) Continuing the search for a law of mortality. *Popul Dev Rev* 22:231-264.

Coles LS. (2004) Demography of human supercentenarians. *J Gerontol A Biol Sci Med Sci* 59(6):B579-586.

Comfort A. (1964) Aging: *The Biology of Senescence*. Holt, Rinehart and Winston, Inc.

Cournil A, Robine, J-M & Vaupel JW. (2005) *Structure and Contents of the International Database on Longevity*. Background paper for the IDL Monograph (To be published).

Curtsinger JW, Fukui, HF, Townsend DR & Vaupel JW. (1992) Demography of genotypes: Failure of limited life-span paradigm in Drosophila Melanogaster. *Science* 258:461-463.

Cutler RG (1985) La biologie du vieillissement et de la longévité. *Gerontologica Biomedica Acta* 1:7-61.

Dépoid F. (1973) La mortalité des grands vieillards. *Population* (4/5):755-792.

Desjardins B & Bourbeau R. (2005) *The Emergence of supercentenarians in Canada* Background paper for the IDL Monograph (To be published).

Fries JF (1980) Aging, Natural Death, and the Compression of Morbidity. *The New England Journal of Medicine* 303(3): 130-135.

Gavrilov LA & Gavrilova NS. (2001) The Reliability Theory of Aging and Longevity. *J Theo Biol* 213:527-545.

Gavrilov LA & Gavrilova NS. (2003) The quest for a general theory of aging and longevity. *Science of Aging Knowledge Environment*. www.sageke.sciencemag.org/cgi/content/full/sageke;2003/28/re5.

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 B Biol Sci* 115:513-585.

Greenwood M & Irwin JO. (1939) The Biostatistics of senility. *Human Biology*. 11 :1-23.

Hayflick L (1981) *Prospects for Human Life Extension by Genetic Manipulation*. pp. 162 – 179 in *Aging: A Challenge to Science and Society*, edited by Danon, D., N. W. Shock, M. Marois. Volumn I. Oxford University Press.

Hayflick L (1996) *How and Why we Age?* New York: Ballantine Books.

Horiuchi S & Wilmoth JR. (1998) Deceleration in the age pattern of mortality at older ages. *Demography* 35(4):391-412.

Kannisto V (1994) *Development of Oldest-Old mortality, 1950-1990: Evidence from 28 developed countries*. Odense Monographs on Population Aging 1, Odense University Press, Odense.

Kannisto V. (1996) *The advancing frontier of survival*. Odense: Odense University Press.

Kannisto V. (1999) Discussion on the paper by Thatcher. *J R Statist Soc* 162: 3.

Kestenbaum BM & Fergusson BR. (2002) *Mortality of the extreme aged in the United States in the 1990s. Based on improved Medicare data*. In *Living to 100 and beyond: Survival at advanced ages*. Lake Buena Vista, Florida, January 17-18.

Kestenbaum BM & Fergusson BR. (2005) *United States*. Background paper for the IDL Monograph (To be published).

Lynch SM & Brown JC. (2001) Reconsidering mortality compression and deceleration: an alternative model of mortality rates. *Demography* 38:79-95.

Maier H & Scholz R. (2005) Age validation of individuals age 105 and above in German. Background paper for the IDL Monograph (To be published).

Makeham WM. (1860) On the law of mortality and the construction of annuity tables. *J Inst Actu* 8:301-10.

Manton KG. (1999) Dynamic paradigms for human mortality and aging. *J Gerontol A Biol Sci Med Sci* 54A(6):B247-254.

McCormack J. (2005) Being very old in a young country: centenarians and supercentenarians in Australia. Background paper for the IDL Monograph (To be published).

Meslé F & Vallin J. (2000) *Comment améliorer la précision des tables de mortalité aux grands âges : le cas de la France ?* – Paris, INED, 26 p. (Communication présentée au Colloque de l'AIDELF "Vivre plus longtemps, avoir moins d'enfants, quelles implications ?", Byblos, Liban).

Meslé F, Vallin J, Robine J-M, Desplanques G & Cournil A. (2005) *Is it possible to measure life expectancy at 110 in France?* Background paper for the IDL Monograph (To be published).

Moivre A. (de) (1725) *Annuities upon lives: or valuation of annuities upon any number of lives*. London.

Olshansky SJ, Carnes BA. (1997) Ever since Gompertz. *Demography* 34(1):1-15.

Perks W. (1932) On some experiments on the graduation of mortality statistics. *Journal of the Institute of Actuaries* 63:12-40.

Poulain M. (2005) *More on the validation of supercentenarians*. Background paper for the IDL Monograph (To be published).

Robine J-M. (2001) A new biodemographic model to explain the trajectory of mortality. *Experimental Gerontology* 36:899-914.

Robine J-M & Vaupel JW. (2002) Emergence of supercentenarians in low-mortality countries. *North American Actuarial Journal* 6(3):54-63.

Robine J-M. (2003) Life course, environmental change, and life span. *Population and Development Review* 29(Suppl):229-38.

Rosenwaike I & Stone LF. (2003) Verification of the ages of supercentenarians in the United States: Results of a matching study. *Demography* 40(4):727-739.

Saito Y. (2005) *Supercentenarians in Japan*. Background paper for the IDL Monograph (To be published).

Skytthe A, Tillo F, Hervonen A, Hoelseth D & Jeune B. (2005) *Supercentenarians in the Nordic Countries*. Background paper for the IDL Monograph (To be published).

Thatcher AR, Kannisto V & Vaupel JW. (1998) *The Force of Mortality at Ages 80 to 120*. Odense: Odense University Press.

Thatcher AR. (1999) The long-term pattern of adult mortality and the highest attained age. *J R Statist Soc* A 162: 5-43.

Thatcher AR, Kannisto V & Andreev K. (2002) The Survivor Ratio Method for Estimating Numbers at High Ages. *Demographic Research*, Vol 6-1. (Available online at www.demographic-research.org).

Vanfleteren JR, De Vreese A & Braeckman P. (1998) Two-parameter logistic and Weibul equations provide better fits to survival data from isogenic population of Caenorhabditis elegans in axenic culture than does the Gompertz model. *J Gerontol A Biol Sci Med Sci* 53: B393-403.

Vaupel, James W (2003) *Post-Darwinian Longevity*. In: JR Carey and S Tuljapurkar (Ed.). *Life span : evolutionary, ecological, and demographic perspectives*. New York: Population Council, 2003: 258-269 (Supplement: Population and Development Review, 29, 2003)

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

Vaupel JW, Johnson TE & Lithgow GJ. (1994) Rates of mortality in populations of Caenorhabditis elegans. *Science* 266:826.

Vaupel JW, Carey JR, Christensen K, Johnson TE, Yashin AI, Holm NV, Iachine IA, Kannisto V, Khazaeli AA, Liedo P, Longo VD, Zeng Y, Manton KG & Curtsinger JW. (1998) Biodemographic trajectories of longevity. *Science* 280:855-860.

Vaupel JW, Carey JR & Christensen K. (2003) Aging. It's never too late. *Science* 301(5640):1679-81.

Vincent P. (1951) La Mortalité des Vieillards. *Population* 6:181-204.

Walford R. (1985) *Maximum Life Span*. New York: WW Norton & Company.

Weitz JS & Fraser HB. (2001) Explaining mortality rate plateaus. *Proc Natl Acad Sci* 98(26):15383-15386.

Table 1: Number of records by level of validation (IDL, version March 2004, persons deceased at age 110 and over, n=677)

Level of validation	Males	Females	Total
0	19	110	129
1	17	130	147
2	36	365	401
Total	72	605	677

Table 2: Number of records by country of death (IDL, version March 2004, persons deceased at age 110 and over, n=660)

Country of death	Males	Females	Total
United States	34	318	352
Japan	13	62	75
France	6	60	66
England and Wales	3	58	61
Italy	5	21	26
Netherlands	4	14	18
Canada	1	16	17
Australia	1	11	12
Sweden	0	9	9
Belgium	0	5	5
Norway	1	4	5
Spain	0	3	3
Denmark	0	2	2
Finland	0	2	2
Austria	1	0	1
Republic Czech	0	1	1
Ireland	0	1	1
Morocco	1	0	1
New Zealand	0	1	1
Romania	0	1	1
South Africa	0	1	1
Total	70	590	660

Table 3: Number of records by dataset (IDL, version March 2004, persons deceased at age 110 and over, n=660)

	Source	Females	Males	Total	Birth range		Death range		Validation level
<u>Dataset 1</u>									
United Sates	SSA	295	31	326	1867	1889	1980	2003	2
France	INSEE	46	3	49	1875	1891	1987	2002	2
England & Wales	ONS	51	2	53	1856	1892	1968	2002	1
Japan	MHW	25	4	29	1884	1890	1996	2001	1
Italy	ISTAT	19	2	21	1863	1889	1973	2000	*
Denmark	CRS	2	0	2	1884	1889	1996	2000	1
Canada / Quebec	ISQ	6	0	6	1872	1889	1983	2001	2
Sub total	-	444	42	486	1856	1892	1968	2003	-
<u>Dataset 2</u>									
Belgium	-	5	0	5	1882	1890	1993	2002	1
Finland	-	2	0	2	1878	1887	1989	2000	1
Netherlands	-	13	3	16	1745	1887	1857	2001	1
Norway	-	0	1	1	1892	1892	2003	2003	1
Sweden	-	1	0	1	1892	1892	2003	2003	1
Sub total	-	21	4	25	1745	1892	1857	2003	-
<u>Dataset 3</u>									
Australia	-	6	1	7	1880	1891	1990	2002	1-2
Canada	-	10	1	11	1877	1892	1988	2002	1-2
France	-	12	2	14	1866	1890	1977	2002	2
Ireland	-	1	0	1	1820	1820	1932	1932	1
Italy	-	0	2	2	1889	1890	2002	2003	1
United States	-	1	0	1	1887	1887	2002	2002	2
Sub total	-	30	6	36	1820	1892	1932	2003	-
<u>Dataset 4</u>									
16 countries	-	95	18	113	1792	1892	1903	2003	0
Total	-	590	70	660	1745	1892	1857	2003	-

* Validation in progress

Figure 1: Number of persons having reached 110 years of age in the year (IDL, version March 2004, persons deceased at age 110 and over, datasets 1, 2 & 3, n=547)

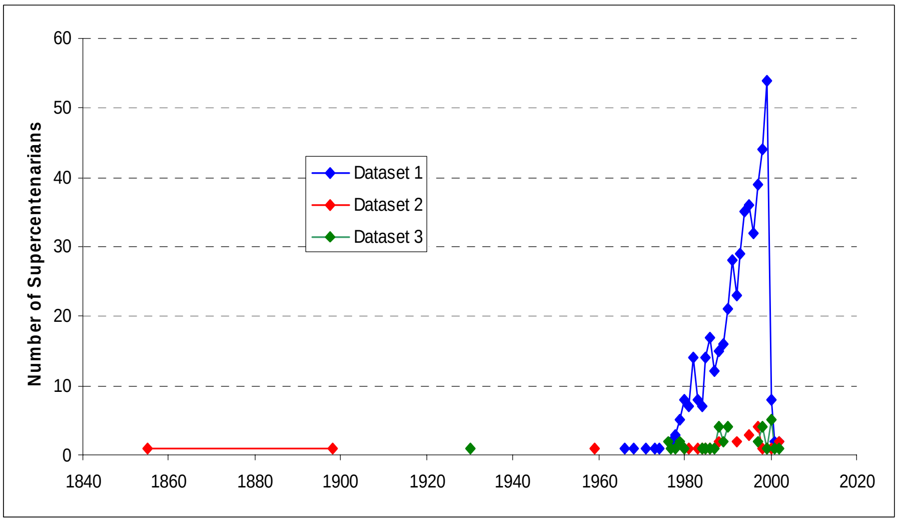

Figure 2: Number of persons having reached 110 years of age in the year (IDL, version March 2004, persons deceased at age 110 and over, dataset 4, remaining records, n=113)

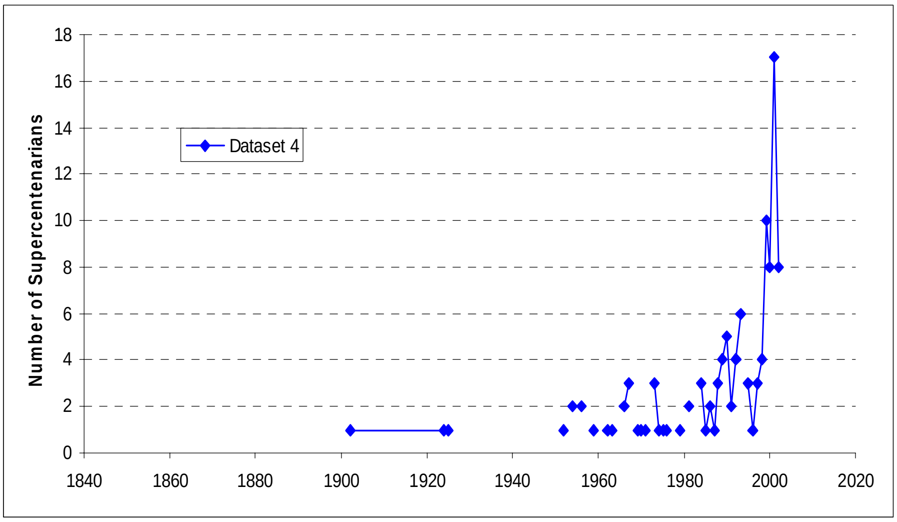

Figure 3: Number of persons having reached 110 years of age in the year (IDL, version March 2004, persons deceased at age 110 and over, dataset 1, institutional sources, n=486)

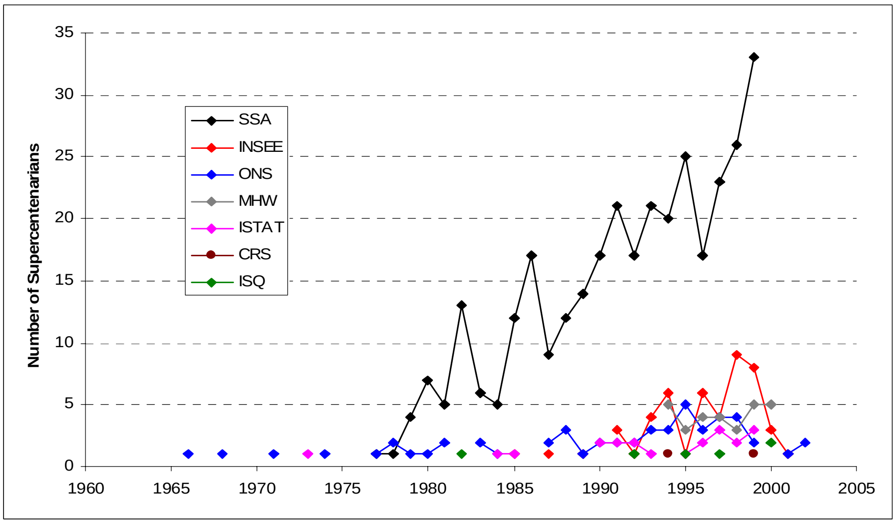

Figure 4: Age structure by dataset (IDL, version March 2004, persons deceased at age 110 and over, n=660)

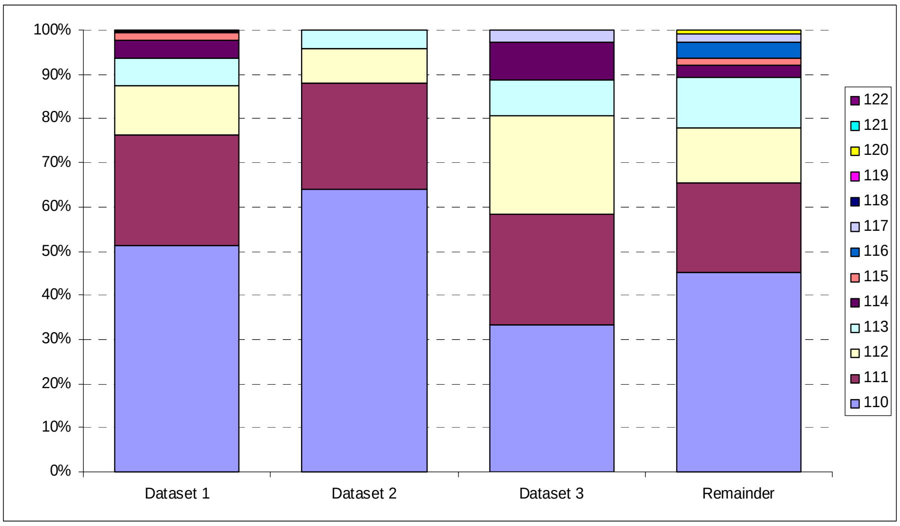

Figure 5: sex-ratio by dataset (IDL, version March 2004, persons deceased at age 110 and over, n=660)

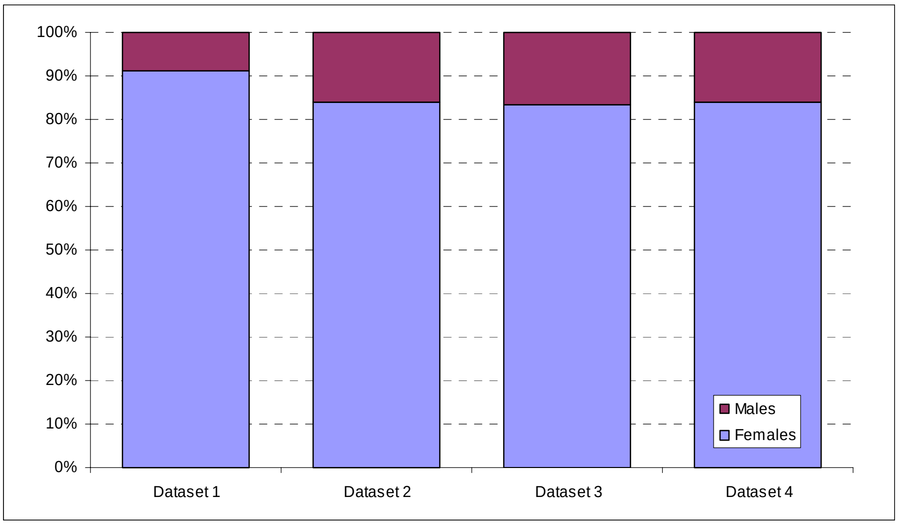

Figure 6: Maximum Reported Age at Death - MRAD (IDL, version March 2004, persons deceased at age 110 and over, dataset 1, institutional sources, n=486)

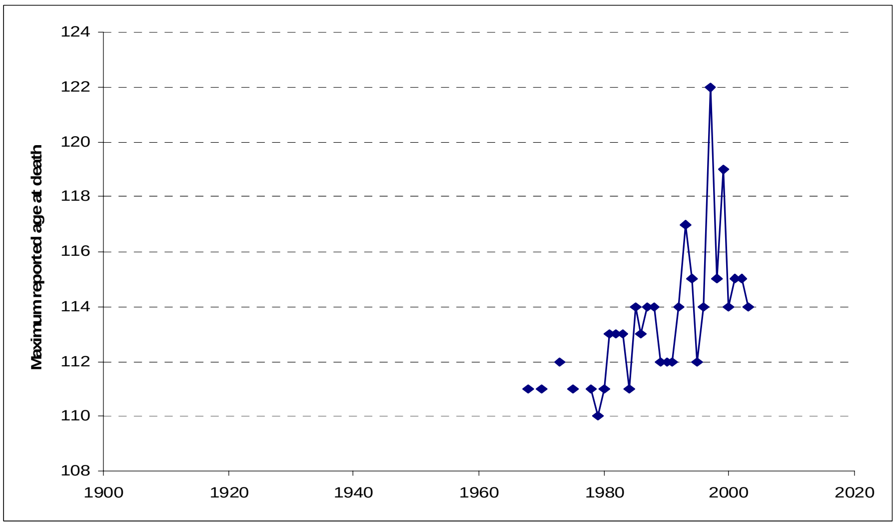

Figure 7: Estimated survival probabilities for female US supercentenarians in extinct cohorts. (Cohorts 1870-80: n=104; 1881-87: n=134).

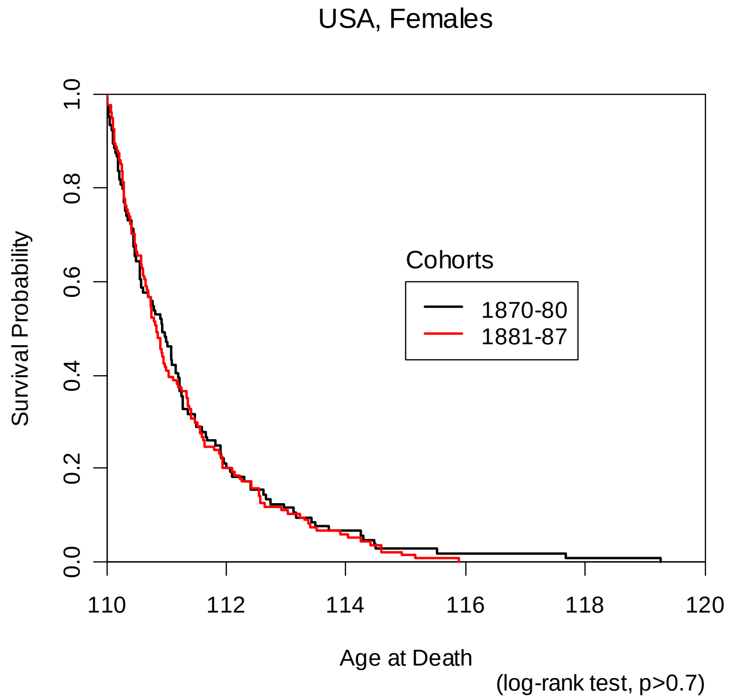

Figure 8: Survival probabilities and integrated hazard for female US Supercentenarians (extinct cohorts only) plus fitted Exponential distribution (maximum likelihood). The confidence level is $(1-\alpha)=0.95$, the estimated parameter of the Exponential distribution is $\hat{\lambda}=0.764$.

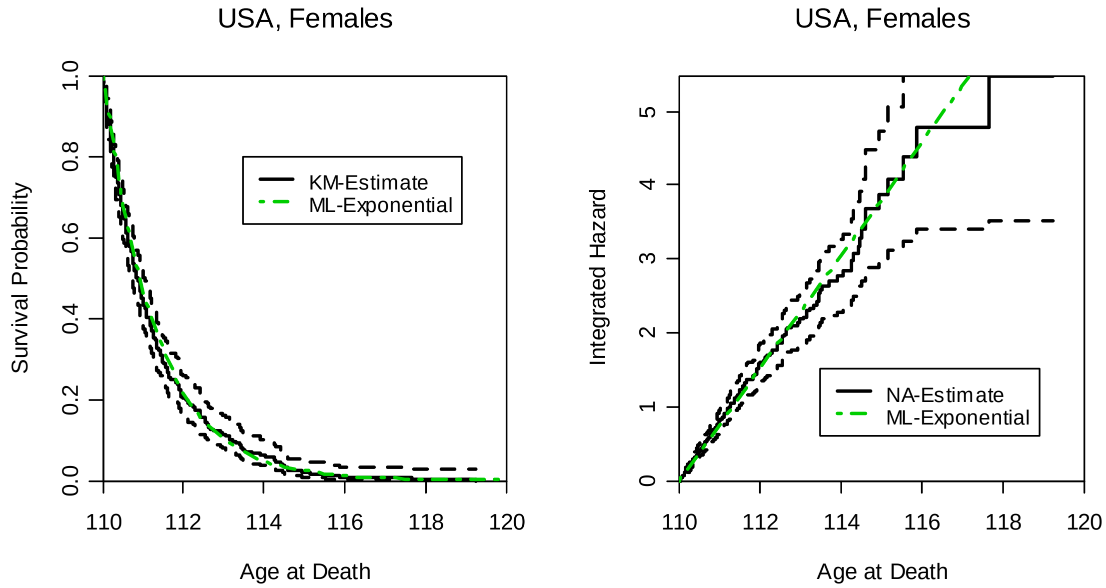

Figure 9: Survival probabilities and integrated hazard for male US supercentenarians (extinct cohorts only) plus fitted Exponential distribution (maximum likelihood). The confidence level is $(1-\alpha)=0.95$, the estimated parameter of the Exponential distribution is $\hat{\lambda}=0.608$.

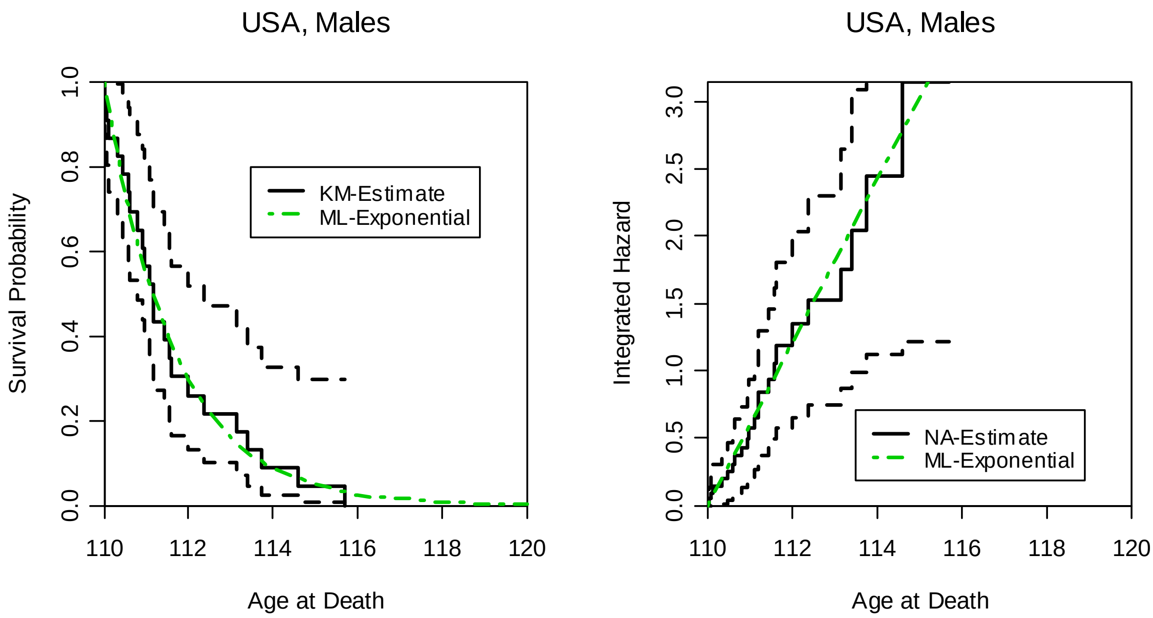

Figure 10a: Survival probabilities and integrated hazard for female supercentenarians in IDL datasets 1 and 2 (extinct cohorts only) plus fitted Exponential distribution (maximum likelihood). The estimated parameter of the Exponential distribution is $\hat{\lambda} = 0.735$.

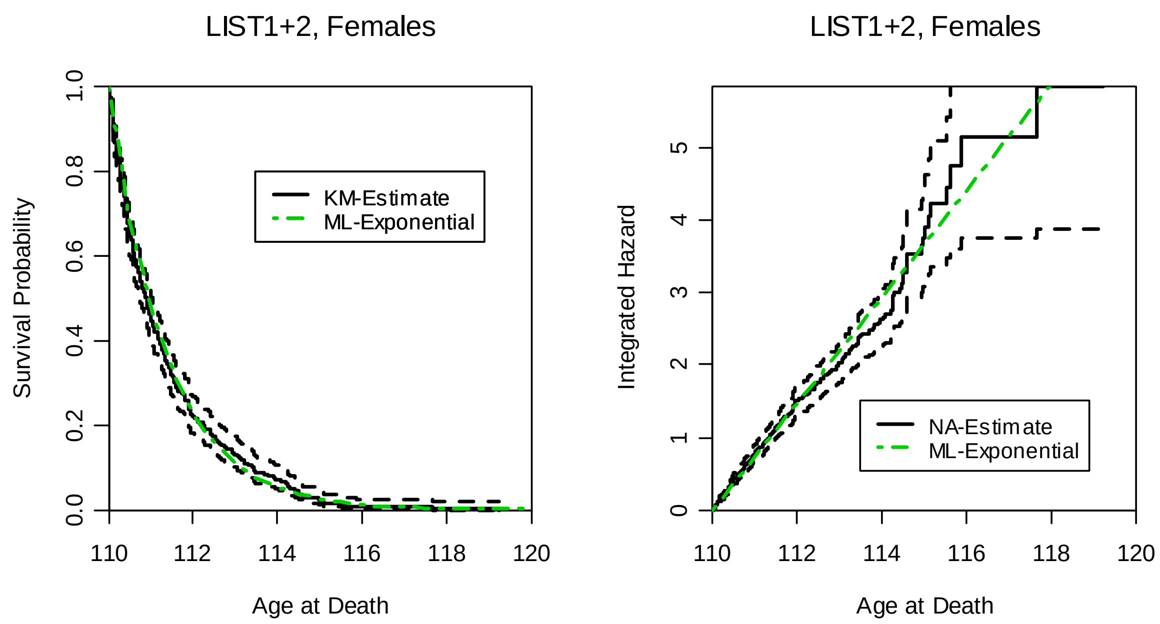

Figure 10b: Survival probabilities and integrated hazard for male supercentenarians in IDL datasets 1 and 2 (extinct cohorts only) plus fitted Exponential distribution (maximum likelihood). The estimated parameter of the Exponential distribution is $\hat{\lambda} = 0.623$.

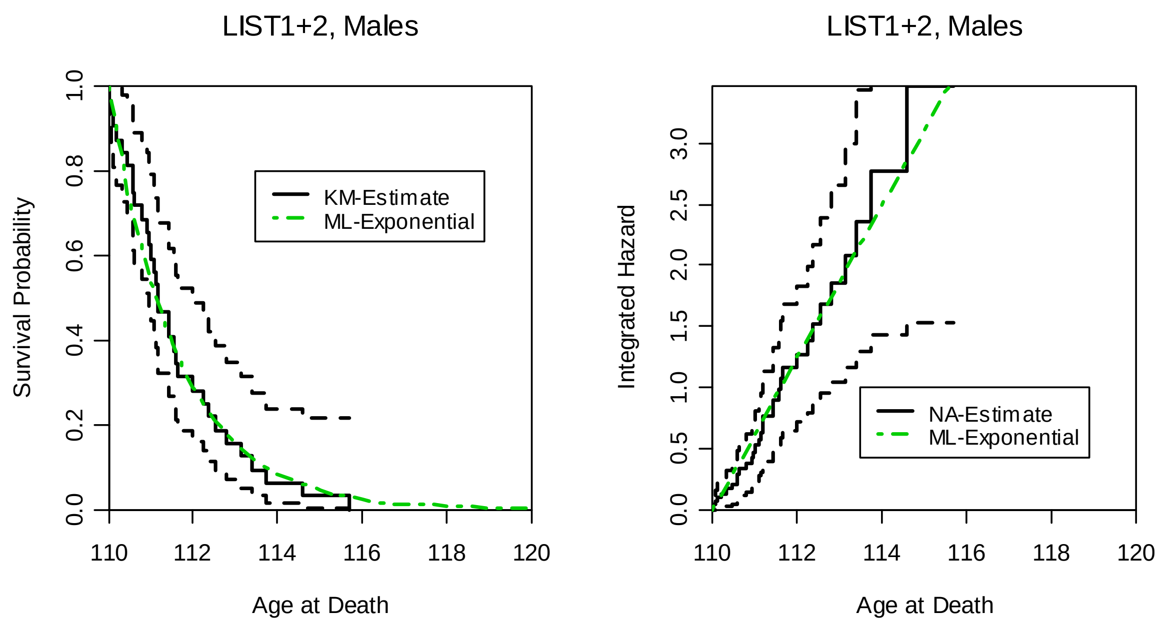

Figure 11a: Survival probabilities and integrated hazard for female supercentenarians in IDL datasets 1 and 2 plus fitted Exponential distribution (maximum likelihood). The estimated parameter of the Exponential distribution is $\hat{\lambda} = 0.744$.

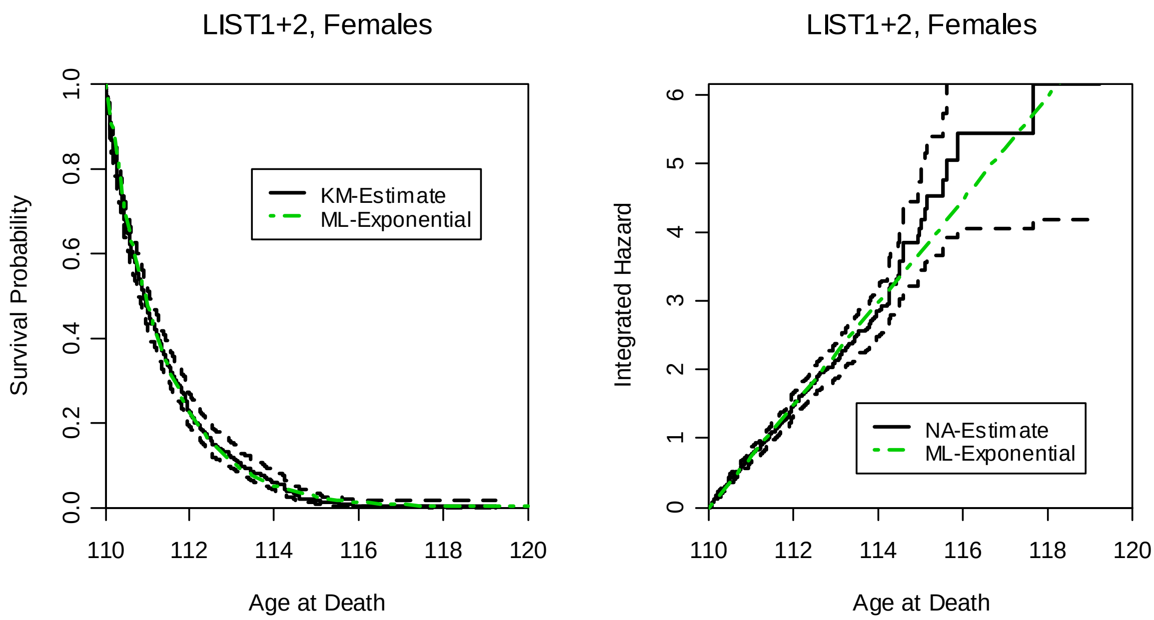

Figure 11b: Survival probabilities and integrated hazard for male supercentenarians in IDL datasets 1 and 2 plus fitted Exponential distribution (maximum likelihood). The estimated parameter of the Exponential distribution is $\hat{\lambda} = 0.664$