TRAJECTORIES OF EXTREME SURVIVAL IN HETEROGENEOUS POPULATIONS

Elisabetta Barbi *et al.*

I.N.E.D | *Population (english edition)*

2003/1 - Vol. 58
pages 43 à 66

ISSN 1634-2941

Trajectories of Extreme Survival in Heterogeneous Populations

Elisabetta BARBI*, Graziella CASELLI*
and Jacques VALLIN**

Because of the scarcity and heterogeneity of observations, the existence of a limit to the human life span remains debatable. Modelling can provide technical solutions, but empirical testing of its validity is based upon disparate individual data. In this article, Elisabetta BARBI, Graziella CASELLI *and Jacques VALLIN advance the debate by proposing a model that allows for the fact that it applies only to one sub-group of the general population, in which frailty is both specific and widely variable between individuals. The distribution of individual frailty, whether acquired or hereditary, is influenced by the history of the cohorts, which modifies the composition of the group, and by the biological aging processes, whose incidence varies between individuals. Using individual French data, the authors show that the length of life at the oldest ages is increasing and find no evidence to refute the hypothesis that there is no clear limit to the human life span.*

Recent gains in life expectancy among the elderly have noticeably contributed to increase average life expectancy in developed countries. The leading protagonists who have ushered in this new phase of the mortality decline, particularly with respect to degenerative causes, are the old and oldest old, who are attaining thresholds that were deemed unthinkable 30 or 40 years ago. In the 1960s an 80-year-old French woman could expect to live a further 6.4 years, in 1970 7.2 years, and by 1997 as much as 9.4 years.

This new phase, which allows an increasing number of elderly persons to postpone their appointment with death, has generated an unprecedented increase in the number of the elderly and especially the oldest old (Vaupel and Jeune, 1995; Vaupel, 1997). As the millennium dawns the elderly, in increasing numbers, are reaching if not surpassing the 100-year threshold to such an extent as to capture the interest of academics from

*Department of Demography, University of Rome "La Sapienza".
**Institut National d'Études Démographiques, Paris.

various disciplines. In France, for instance, while at the beginning of the fifties there were only 200 centenarians, their number has been increasing so fast since then that it reached almost 3,800 in the 1990 census, and will probably exceed 10,000 in 2002[1].

Demographers too have begun to focus more frequently on longevity. Whether or not there are biological limits to the human life span is a topical issue (Bourgeois-Pichat, 1952; Manton et al., 1991; Vallin, 1993; Thatcher, 1999). Are recent gains due to increased longevity among a growing proportion of the population, an outcome that will speed up the rectangularization of the survival curve? Or are such gains the harbinger of new frontiers that may announce the further "extension" of the survival curve?

Limits to the human life span remain an open debate. Any discussion on the limits of maximum human survival, whether they in fact exist, and how they may be manifested, is misleading unless the impact of population heterogeneity on mortality is also considered. Each individual in the population is endowed with a specific genetic patrimony that grants a certain kind of resistance to death and disease within the course of life. Moreover, during their life, individuals undergo different experiences which may further diversify their ability "to withstand destruction" (Gompertz, 1825, p. 517).

Demographers for some time have observed that the speed of increase of the force of mortality slows down at older ages, even though heretofore the poor quality of the data at their disposal undermined the results obtained (Vaupel, 1997). Finally, highly reliable mortality data for the oldest old are available for developed countries, suggesting that the human mortality trajectory for these ages does not follow Gompertz's exponential function (Horiuchi and Coale, 1990; Kannisto, 1994, 1996; Kannisto et al., 1994; Thatcher et al., 1998; Wilmoth and Horiuchi, 1999). One explanation may be that individuals of the same generation may be vulnerable to death in different ways (Vaupel et al., 1979; Vaupel and Yashin, 1985; Vaupel and Carey, 1993). Thus the observed deceleration in the force of mortality may be traced to changes in the composition of heterogeneous populations.

From one generation to the next, possible modifications in the heterogeneity of individual frailty may provoke changes in the population's survival function. In turn, these changes may open new horizons in survival. The population's (observed) survival trends may be related to different (unobserved) individual processes. Different hypotheses regarding the heterogeneity of individual frailty may generate the same population mor-

[1] At the 1999 census, 11,593 centenarians were enumerated. However, contrary to the previous census, which devoted particular attention to the oldest ages and checked the returns, this one obviously overestimated the number of centenarians. According to INSEE estimates, this number was 6,840 on 1 January 1998, which means that it grows at a rate of 8% per year and will probably exceed 10,000 before the end of 2002 (Vallin and Meslé, 2001b).

tality model. If there is a decline in mortality at the population level, the dynamics of heterogeneity in the change-over from one generation to the next could imply an increase in the proportion of individuals who, owing to genetic (hereditary) or environmental (acquired) factors, are less vulnerable to death and disease. What would then occur would be a process of homogenization among individuals favourable to increasingly lower levels of frailty. Heterogeneity, therefore, would only involve a different composition of the population which, on average, would enjoy lower mortality even though none of the individuals in that population experienced lower mortality.

On the other hand, the observed mortality curve for the population might be the outcome of new gains in survival by its more robust individuals who would thus experience more favourable mortality trajectories. If this was the case, the population would remain markedly heterogeneous. Heterogeneity would not involve changes in the composition of the population, but rather changes in individual mortality trajectories.

Most likely the population's mortality curve may be generated by both processes, that is jointly by changes in the composition of the population and by further reductions in individual mortality trajectories.

The importance of distinguishing unobservable processes resides in the fact that these may have different implications for longevity and the existence of a limited life span. If the observed mortality curve is merely the result of a process of homogenization of human frailty, the upper age at death could remain unchanged over time and thus confirm the existence of a final threshold beyond which it is impossible to aspire. If the population's mortality curve depends on favourable changes in individual mortality trajectories, the decline in mortality among robust individuals could shift the upper age at death still further. Should this process be an ongoing one, the theory of a pre-established limit to the human life span would no longer hold.

With scenarios such as these, the analysis of survival should take into consideration the impact of the population's heterogeneity on individual frailty. The underlying problem is that frailty cannot be observed and mathematical models are required that include it as a variable to be estimated. These *frailty models*, as they are known, are part of the family of survival analyses. The underlying concept of frailty is based on the hypothesis that individuals surviving into old age enjoy a more favourable force of mortality during their lives, but that the rate of increase with age is similar for all those belonging to the same generation. It has been shown recently that it is not sufficient to consider the heterogeneity of the traditional models, known as *level heterogeneity*, but that the heterogeneity in the speed of mortality increase with age, that is *slope heterogeneity,* may also be involved (Barbi, 1999; Wilmoth and Horiuchi, 1999; Yashin et al., 2001). While the first source of heterogeneity in frailty is due to (*proportional*) individual differences in the level of mortality, the second source

of variability is given by (*non-proportional*) individual differences in the rate of aging. Both kinds of individual differences are equally plausible theoretically and should be taken into account.

To analyse variations in the rate of individual mortality increase with age and to investigate the implications they may have for the mortality of a population, a *mixture frailty model* is applied which incorporates both sources of heterogeneity. The mixture frailty model and the classic frailty model that accounts only for level heterogeneity, have been applied to the mortality data of French female cohorts born between 1820 and 1879. The survival trajectories obtained with these models have been used to estimate the maximum life span. Differences in the speed of aging may play a crucial role in making this estimate. A parametric approach has been applied, as suggested by the classic theory of extreme values (Gumbel, 1937).

A further application involves French centenarians (INSEE data base, 1997[(2)]). Since the data are individual in this instance, a non-parametric approach (Aarssen and de Haan, 1994) will be used to estimate the maximum life span for the cohort born between 1870 and 1879.

Before commenting on the results obtained, the following sections offer a brief description of the models and data used.

I. From classic frailty models to a mixture model

The frailty model, introduced by James Vaupel, Kenneth Manton and Eric Stallard in 1979, is a statistical model that accounts for heterogeneity in individual frailty. Let

$$\mu(x, z) = z\mu(x, 1)$$

be the force of mortality for an individual aged x with a frailty equal to z where $\mu(x, 1)$, the *baseline function*, is the force of mortality for a standard individual of the same age x whose frailty is set equal to 1.

If z at birth is distributed as a Gamma probability distribution:

$$f(z) = \lambda^k z^{k-1} e^{-\lambda z} / \Gamma(k)$$

(2) We are grateful to the Institut National de la Statistique et des Études Économiques (INSEE), Paris, for supplying individual mortality data used in accordance with the Data Protection Legislation.

with the mean set equal to 1 and variance σ^2, then the population's force of mortality at age x may be expressed as:

$$\bar{\mu}(x) = \mu(x, 1)\bar{s}(x)^{\sigma^2}$$

where $\bar{s}(x)$ are survivors from birth to age x.

Notice that the classic frailty model assumes, as is often the case in survival analysis, the proportionality of mortality risks. In other words, differences are hypothesized in individual mortality risks only with regard to the level of mortality (*level heterogeneity*).

The mixture frailty model described below (Barbi, 1999) takes also into consideration the *slope heterogeneity*, that is individual differences in the speed of increase in mortality with age.

As before, we assume that the proportional frailty z is Gamma distributed with mean 1 and variance σ^2. Moreover, it is assumed that the distribution of the non-proportional frailty τ at birth is discrete with a probability distribution

$$\Pr(\tau_i) = \pi_i \geq 0, \sum_{i=1}^{n} \pi_i = 1$$

Finally, we assume that the distributions of z and τ are independent.

Let

$$\mu(x, z, \tau_i) = z\mu(x, 1, \tau_i)$$

be the force of mortality given frailties z and τ_i where $\mu(x, 1, \tau_i)$, the *baseline hazard function* of the ith population group, is the standard individual's force of mortality, with regard to the level heterogeneity, of the population group with frailty τ_i, with respect to the slope heterogeneity in frailty.

In other words, $\mu(x, z, \tau_i)$ expresses the individual's force of mortality at age x, with frailty z with regard to the level of mortality, and frailty τ_i with respect to the rate of mortality increase with age. It may be noted that both frailty definitions assume that individuals are born with set frailty levels z and τ_i, and that these remain so throughout their lives.

Therefore, according to these hypotheses, the population's force of mortality is the weighted average:

$$\bar{\mu}(x) = \sum_{i=1}^{n} \pi_i(x)\mu(x, 1, \tau_i)s_{x|\tau_i}(x, \tau_i)^{\sigma^2}$$

where $s_{x|\tau_i}(x, \tau_i)$ may be considered as the mean survival function with respect to the level heterogeneity, of individuals belonging to the *i*th group, and $\pi_i(x)$ are positive weights, denoting the proportion of each group, with

$$\sum_{i=1}^{n} \pi_i(x) = 1$$

In other words, it is assumed that the population is sub-divided into *n* sub-groups defined by *n* different rates of mortality increase with age and that individuals within these subgroups are homogeneous with regard to the rate of mortality increase, but heterogeneous with regard to the level of mortality. The degree of level heterogeneity shared by all the *n* subgroups, is given by the variance of the Gamma distribution assumed to describe the proportional frailty *z*.

In the applications performed here, it is assumed that the population is only divided into two subgroups and that the standard force of mortality with respect to both frailties $\mu(x, 1, 1)$ is described by a Gompertz-Makeham function with parameters *a*, *b* and *c*. It is assumed that the baseline hazard function of the *i*th group is expressed by:

$$\mu(x, 1, \tau_i) = a\exp(\tau_i bx) + c$$

or rather,

$$\mu(x, 1, \tau_i) = a\exp(b_i x) + c$$

The model was estimated using a maximum likelihood method. The likelihood function is maximized by an iterative numerical procedure. The inverse of the observed information matrix yields an asymptotic estimate of the variance and covariance matrix of the estimated parameters.

II. The upper age limit at death: parametric and non-parametric approaches

1. Parametric approach

The mortality curves obtained by applying the classic frailty model and the mixture model were used to estimate the maximum life span of the birth cohorts considered. The approach followed was the standard extreme value theory. Perhaps it is worth calling to mind a few fundamental steps. Let us consider a cohort of individuals, assuming that N members survive at age x_0. When all these N individuals die, this will yield N ages at death, of which one will be the oldest age reached by that cohort. As Roger Thatcher (1999) points out, this age is not a potential value. What this does represent is the upper age at death of a sample of size N that experienced the same risks as the cohort studied. Obviously, the positive or negative experiences of a cohort do not only affect survival, but also determine its maximum life span. Thus the oldest age reached by members of the cohort will have a probability distribution derived from number N and the estimated force of mortality.

Let:

N be the number of members of a birth cohort who reach age x_0 with $N = c(x_0)$;

ω_N be the oldest age reached by N individuals of the cohort;

$c(x)$ be the number of members of the cohort who reach age x among those who reach age x_0, that is among N members:

$$c(x) = Ns(x) = N\exp\left(-\int_{x_0}^{x}\mu(t)dt\right)$$

where $s(x)$ is the probability that a person will survive from age x_0 to a given age x.

The probability that the maximum life span will be less than age x will be equal to:

$$\Pr(\omega_N < x) = \left(1 - \frac{c(x)}{N}\right)^N \cong e^{-c(x)}$$

if N is large.

With these two formulas, the probability distribution of ω_N can be calculated.

It may be shown (Fisher and Tippett, 1928) that under certain circumstances and when N tends towards infinity, the probability distribution of ω_N tends towards a limiting form. This outcome was generalized and used by Gumbel (1937) in his study of mortality. He showed that when N is large, the probability distribution of ω_N is close to the value of x which satisfies:

$$c(x) = 1$$

or rather

$$s(x) = 1/N$$

The age which satisfies this equation is obtained by interpolation, while the percentiles were obtained using the exact distribution of the maximum age at death.

2. Non-parametric approach

As individual data were available it was possible to exploit more recent methodological advances in the extreme values theory (Aarssen and de Haan, 1994). Here we recall only some fundamental steps, and methodological details may be taken from the study cited.

Let X_1, X_2,... be independent random variables, all with the same probability distribution, and F the common distribution function. Suppose that the normalized largest value of a sample of size n converges in distribution to a proper probability distribution, i.e. there exist sequences of constants $a_n > 0$ and b_n such that:

$$P\left\{\frac{\max(X_{1,\ldots,}X_n) - b_n}{a_n} \leq x\right\} = F^n(a_n x + b_n)$$

converges for all x to some limit distribution function $G(x)$ which assumes at least 3 values. The function G can only be one out of a limited class of functions. A very useful way to describe this class is:

$$G(x) = G_\lambda(x) = \exp -(1 + \lambda x)^{-1/\lambda}$$

for some $\lambda \in \mathbf{R}$ and for all values of x for which $1 + \lambda x > 0$. In fact this representation describes the functions G as a continuous class depending on one parameter. Therefore we have:

$$\lim_{n \to \infty} F^n(a_n x + b_n) = G_\lambda(x)$$

for some $\lambda \in \mathbf{R}$, and a choice of $a_n > 0$ and b_n.

For fixed $\lambda \in \mathbf{R}$, the question then is for what functions F this relation holds. This set of distribution functions F is called the domain of attraction of $G_\lambda(x)$ for this particular λ. If the relation holds for a distribution function F with $\lambda > 0$, then the right end point of the distribution F is infinite. If this form holds for a distribution function F with $\lambda < 0$, then the right end point of F is finite. Hence the analysis is performed in two steps:

a) Test the hypothesis: $\lambda \geq 0$.

b) If the hypothesis is rejected, estimate the right end point (x^*) of F representing the life span distribution of our population.

The estimators λ and x^* are functions of the k upper order statistics, that is, the number of the largest observations.

III. Data

The present analysis is based on the mortality experienced by French female cohorts born from 1820 to 1879. Cohort life tables have been computed from death rates (Vallin and Meslé, 2001a) up to the age of 85 or 90 (depending on the quality of the data) but above these ages the probability of dying was calculated directly from deaths classified by age and year of birth according to Vincent's (1951) extinct generations method[3]. Above age 103, particularly for cohorts born at the beginning of the nineteenth century, the observed values of the probability of dying are based on small numbers and the data are less reliable.

Figure 1 displays trends in life expectancies at ages 70, 80, 90 and 100. It is clear that there was an acceleration in the increase of life expectancy at older ages.

[3] Cohort probabilities of dying used here have been computed for this paper on the basis of Vallin and Meslé's rates and, above age 85, according to Vincent's method. A more complete computation (Meslé and Vallin, 2002) has been used for the French life tables published recently by INED (Vallin and Meslé, 2001a). There are only negligible differences between the two sets.

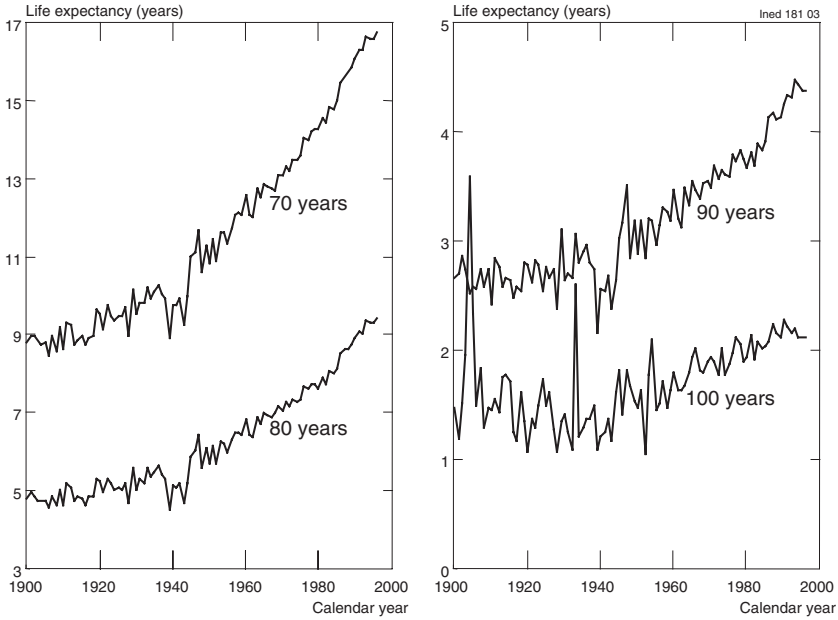

Figure 1.– Changes in life expectancies at 70, 80, 90 and 100 years, computed in French female cohorts born from 1815 to 1884 by the extinct generations method

Source: Vallin and Meslé, 2001.

The starting age (60 years) selected for the model is by no means a random choice. As pointed out above, the drawback of frailty models is that they assume each individual to be born with a personal frailty level that remains fixed for life. This is somewhat restrictive as it fails to reflect reality. Individuals during the course of their life undergo different experiences that may either positively or negatively affect their life potential and thus their vulnerability to death and disease. By choosing 60 as the age to start the observation it is implicitly assumed that frailty is fixed for that age onwards and not from birth. The probability distribution to describe the heterogeneity of the population thus concerns the heterogeneity in frailty for the 60-year-olds. Thus we can capture both initial (genetic) heterogeneity as well as that acquired and accumulated to that age. This would appear to be a more realistic proposition. Ivan Iachine's results obtained applying a frailty model to Danish twins (see Vaupel et al., 1998) suggest that half the individual differences in life expectancy after 30 may be traced to everything that happened before that age, i.e. to genetic factors as well as to those acquired between birth and age 30. About a third of

this 50% would be due to genetic factors and the remaining two thirds to environmental or external factors, such as the individual's socio-economic, nutritional and health status. The same author argues that among a group of persons aged 30 with an average duration of life of 60 years, 80% of the differences in life span may be attributed to the different risk factors accumulated until then.

The INSEE individual mortality database is used for the non-parametric estimation of the maximum life span. These data regard extinct cohorts of French women aged 98 years and over, born between 1870 and 1879. The 10 cohorts chosen include the exceptional "super-centenarian" Jeanne Calment, who was born in 1875 (on 21 February) and died in 1997 (on 4 August) at 122.5 years (Allard et al., 1994, 1998). This data set totals 16,122 persons, a number deemed suitable for the non-parametric approach.

IV. Results and discussion

The two frailty models described in Section 1 and the two approaches to the estimation of the maximum life span shown in Section 2 have been applied to French data. Data for French women born in 1820 to 1879 were divided into 6 groups, each consisting of 10 cohorts. Table 1 displays life expectancy at ages 70, 80, 90 and 100 years for each group.

TABLE 1.– EXPECTATION OF LIFE AT 70, 80, 90 AND 100 YEARS FRENCH FEMALES BORN FROM 1820-29 TO 1870-79

Age	Groups of cohorts					
	1820-29	1830-39	1840-49	1850-59	1860-69	1870-79
70	8.64	8.87	9.18	9.66	10.04	11.17
80	4.71	4.96	5.08	5.24	5.66	6.37
90	2.68	2.68	2.69	2.88	3.16	3.42
100	1.37	1.34	1.52	1.50	1.71	1.89

Source: Vallin and Meslé, 2001.

Let us first look at the results of the classic frailty model, reported in Table 2 and Figure 2, and then at those for the mixture frailty model, in Table 3 and Figure 2. Comments will then be provided on results regarding the maximum life span for each group of cohorts, using the mortality curves obtained with both models. Finally, the results obtained for the upper age at death by using individual data and a non-parametric approach will also be discussed.

1. The role of heterogeneity

As Table 2 shows, the level of heterogeneity (σ^2) using the classic frailty model shows no particular trend, and the values are not high for any of the groups of cohorts, ranging from a minimum of 0.08 for the cohorts born in 1840-49 to a maximum of 0.19 for those born in 1860-69. The latter figure may be attributed to the poor fit of the model to the data (Figure 2e), rather than to a real increase in heterogeneity for these cohorts.

TABLE 2.– CLASSIC FRAILTY MODEL*
PARAMETERS WITH STANDARD ERRORS

Cohorts	*a*	S.E.	*b*	S.E.	*c*	S.E.	σ^2	S.E.
1820-29	0.01825	0.00029	0.10938	0.00091	0.00769	0.00035	0.11737	0.00440
1830-39	0.01744	0.00026	0.11000	0.00087	0.00670	0.00032	0.13056	0.00444
1840-49	0.01800	0.00025	0.10329	0.00079	0.00478	0.00031	0.08215	0.00382
1850-59	0.01149	0.00014	0.12264	0.00069	0.01040	0.00020	0.15780	0.00365
1860-69	0.01260	0.00015	0.11763	0.00068	0.00651	0.00020	0.19183	0.00398
1870-79	0.01125	0.00013	0.10650	0.00059	0.00750	0.00018	0.09446	0.00324

* $\bar{\mu}(x) = a\exp(bx)s(x)^{\sigma^2} + c$
Source: Vallin and Meslé, 2001.

Results for the mixture frailty model (Table 3) with regard to the level of heterogeneity, broadly confirm those obtained using the curves estimated with the other model. No particular trend emerges, and values are low, the highest being 0.15 for the same group of cohorts (1860-69).

To analyse the heterogeneity in the rate of mortality increase with age (*slope heterogeneity*), reference must be made to the proportion π consisting of the frailer population subgroup defined by the exponential coefficient b_1 and to the proportion $(1-\pi)$ consisting of the more robust population subgroup defined by the coefficient b_2 (Table 3). It clearly emerges that both subgroups are quite distinct for all the cohorts considered, apart from those born in 1870-1879 where the estimated π parameter was not significant. Moreover, two clearly distinct situations may be pinpointed: that for the 1820-29 and 1850-59 cohorts and for the 1830-39, 1840-49 and 1860-69 cohorts.

For the 1820-29 and 1850-59 cohorts, the frailest individuals at 60 years of age comprise more than 70% of the total (71% for the former and 74% for the latter group) and both exponential coefficients, while different (b_1=0.10 and b_2=0.08 for the 1820-29 cohorts and b_1=0.11 and b_2=0.09 for the 1850-59 cohorts), generate an obvious divergence in the speed of mortality increase only for the upper ages. Figures 3a and 3d show how the three mortality curves (for the frailest women μ_1, the most robust women μ_2, and the total population μ) remain distinct until 95 years. After this the

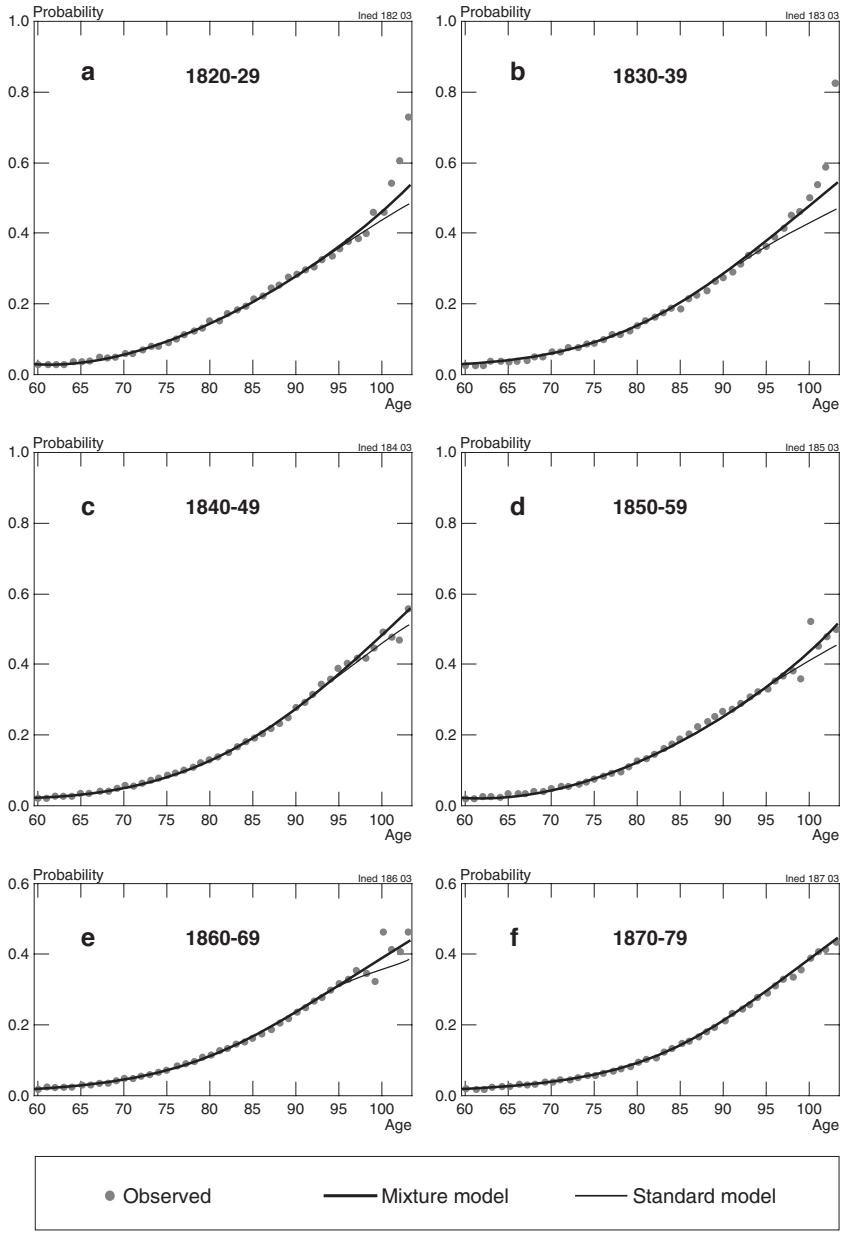

Figure 2.– Observed and estimated probabilities of dying
French female cohorts

Source: Vallin and Meslé, 2001

TABLE 3.– MIXTURE FRAILTY MODEL*
PARAMETERS WITH STANDARD ERRORS

Cohorts	a	S.E.	c	S.E.	b_1	S.E.	π	S.E.	b_2	S.E.	σ^2	S.E.
1820-29	0.02125	0.00030	0.00450	0.00038	0.10156	0.00095	0.71412	0.01938	0.08246	0.00072	1.5E-07	2.85E-08
1830-39	0.01314	0.00028	0.01183	0.00035	0.16087	0.00263	0.22326	0.00953	0.11172	0.00125	0.08201	0.00634
1840-49	0.01636	0.00034	0.00668	0.00040	0.14185	0.00464	0.09495	0.02015	0.10383	0.00124	0.06654	0.00678
1850-59	0.01376	0.00015	0.00789	0.00020	0.11342	0.00068	0.73843	0.01125	0.09131	0.00061	0.00477	0.00242
1860-69	0.00826	0.00015	0.01194	0.00020	0.18211	0.00184	0.23524	0.00552	0.12390	0.00089	0.15168	0.00488
1870-79	0.01126	0.00014	0.00749	0.00019	0.11063	0.00299	0.09955	0.17056	0.10590	0.00160	0.09250	0.00667

* $\bar{\mu}(x) = \sum_{i=1}^{2} \pi_i(x) a \exp(b_i x) s_{x|\tau_i}(x, \tau_i)^{\sigma^2} + c$

Source: Vallin and Meslé, 2001.

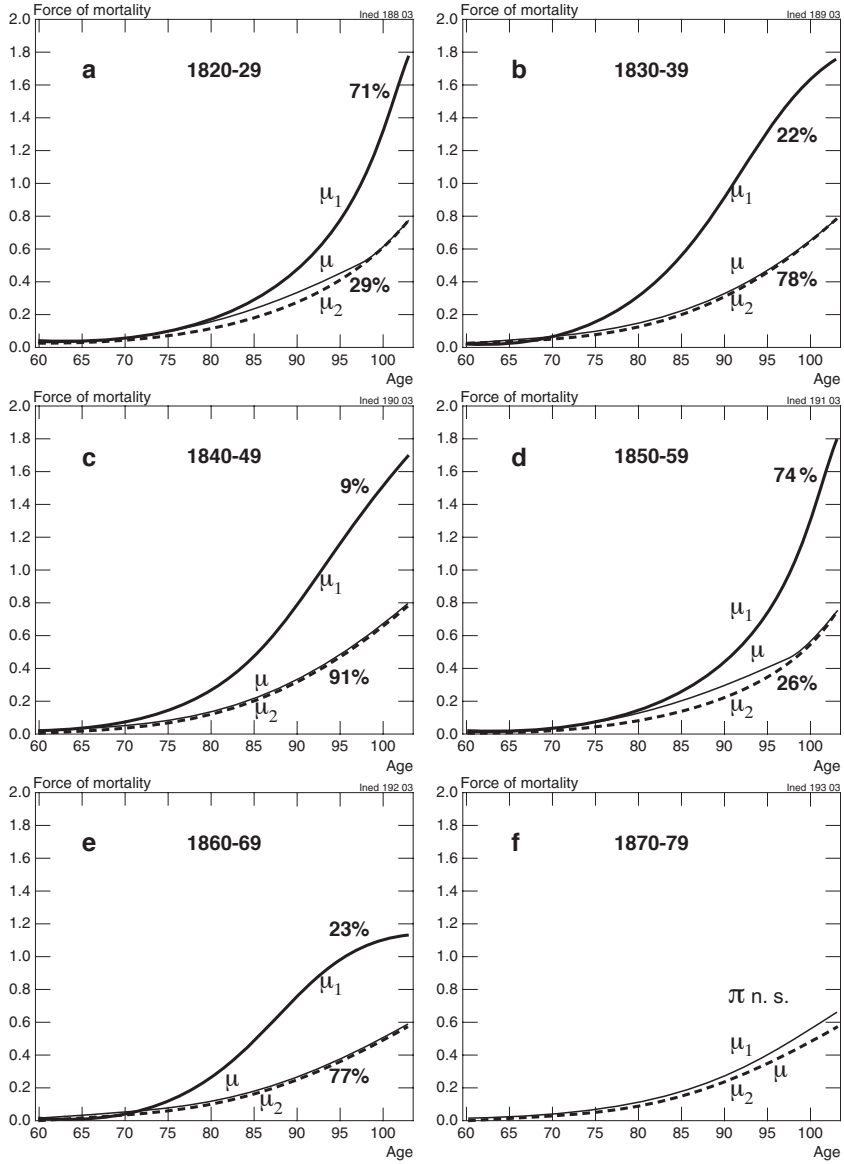

Figure 3.– Estimated force of mortality for subgroups (μ_1 for frailer individuals and μ_2 for stronger individuals), and total population (μ)
French female cohorts

Source: Vallin and Meslé, 2001.

mortality curve for the most robust (about 30% at 60 years) converges and, finally, overlaps with that for the total population. Thus, by the time they reach 100 years of age, the frailer members have died out and most of the population is composed of the more robust female subgroup.

For the 1830-39, 1840-49 and 1860-69 cohorts, the first difference that emerges in comparison with the cohorts discussed before is the decline in the number of 60-year-olds among the frailer groups (22%, 9% and 23%, respectively). Their disadvantage, however, is relevant this time in that soon after reaching 70 their mortality curve and that for the stronger women diverge visibly. This implies that there is a rapid fall in their number and subsequently the mortality curve for the population may not be distinguished from that of the stronger subgroup. The robust women, as they approach 80, represent most of the population. Thus for these cohorts, the youngest old are homogeneous with respect to slope heterogeneity.

A separate comment is warranted for the 1870-1879 cohorts that include Jeanne Calment who lived until the age of 122 and a half. In this instance, as stated above, the parameter indicating the proportion of frailer individuals is not significant. The estimates for the two exponential coefficients b_1 and b_2 are very close. Thus, slope heterogeneity plays no role at the population level and Figure 3f shows how the three curves overlap almost completely. Level heterogeneity for these cohorts, as already seen above, is also quite low. It is worth pointing out, although it may only be a coincidence, that the oldest ages at death were reached by women of these cohorts.

The likelihood-ratio test shows how the mixture frailty model fits better than the classic frailty model for all the groups of cohorts except the last (Table 4). As we just saw, the slope heterogeneity for this group is not significant. However, the high levels reached by the log likelihoods make the test not very reliable.

TABLE 4.– RESULTS OF THE LIKELIHOOD-RATIO TEST

Cohorts	–Log likelihood		Likelihood-ratio test	
	Model 1, df = 4	Model 2, df = 6	χ^2	p
1820-29	5,534,564	5,534,502	124	0.00
1830-39	6,324,447	6,324,158	578	0.00
1840-49	6,659,097	6,659,060	74	0.00
1850-59	6,833,626	6,833,502	248	0.00
1860-69	7,550,707	7,550,073	1,268	0.00
1870-79	7,768,312	7,768,311	2	0.66

Model 1: Classic frailty model.
Model 2: Mixture frailty model.
Source: Vallin and Meslé, 2001.

2. The maximum life span

Parametric approach

Following a parametric approach, the mortality curves obtained with the two frailty models were used to estimate the maximum life span.

According to Table 5, the upper age at death, moving from older to younger cohorts, rises steadily from 108.4 years to 112.5 years using the classic model and from 107.1 years to 112.5 years using the mixture model. Usually figures are higher for the former, particularly for the 1860-1869 cohorts for whom the highest level of heterogeneity was estimated (equal to 0.19), a heterogeneity which caused the speed of mortality increase with age to slow down, excessively so.

Regarding the last group of cohorts, naturally identical results were obtained with both models. Of interest is the fact that for these the highest modal age at death is predicted and the upper percentile (117.7) closely mirrors that observed in real life (117.2 years). However, Jeanne Calment's age at death considerably exceeds the estimated one, and thus is not covered by our models, as is the case with other recent applications (Thatcher et al., 1998).

A more in-depth analysis of the results obtained using the mortality curves estimated according to the mixture model reveals that more substantial gains were obtained in the shift from older to more recent cohorts than with the classic model: 5.4 years in modal age, compared with 4.1 years if the 1860-69 cohorts are excluded due to poor fit. This gain is all the more spectacular when the upper percentiles are considered: 6.9 years versus 3.4 years.

If the upper age limits are analysed taking trends in slope heterogeneity into account, it would appear that this plays a perceptible role in establishing the maximum life span. Estimates obtained depend on the number of individuals *N*, on the estimated mortality level, and the rate of mortality increase with age (or rather exponential coefficients of the Gompertz-Makeham function). The slowing down of this increase, observed in the passage from older to more recent cohorts, not to mention the decline in the force of mortality (Table 3, Figure 3), does not suffice to raise the upper age limit at death. For this to become a reality, increasing numbers of the population would have to experience similarly favourable levels. Data for Table 5 show that the largest gains in the maximum life span occur for those cohorts with the highest degree of homogeneity and the smallest number of fragile individuals. Focusing on the upper percentiles, a jump occurs from the 110.8 years estimated for the first group of cohorts to 112.2 years estimated for the following three groups. The most striking increase is observed in the passage from the 1850-59 to the most recent cohorts (112.1 years to 117.7 years), mirroring the shift from greater heterogeneity to pronounced homogeneity.

TABLE 5.– MAXIMUM LIFE SPAN FOR FRENCH FEMALE COHORTS

Cohorts	Distribution of the highest age						The highest observed ages	
	Classic frailty model			Mixture frailty model			INSEE data	Jeanne Calment's age at death
	Mode	Percentiles		Mode	Percentiles			
		1%	99%		1%	99%	1997	1997
1820-29	108.4	106.4	114.3	107.1	105.6	110.8	104	–
1830-39	109.9	107.8	116.2	107.5	105.8	112.2	111	–
1840-49	108.4	106.6	113.3	107.7	106.0	112.2	107	–
1850-59	110.8	108.7	117.2	108.5	107.2	112.1	109	–
1860-69	115.1	112.5	123.1	112.2	110.1	118.4	109	–
1870-79	112.5	110.7	117.7	112.5	110.7	117.7	117	122.5*

* In the official INSEE files on which the database was built, deaths are not supposed to occur after age 119. Deaths occurring after that age have been randomly redistributed. However, it is possible to get Jeanne Calment's death from other INSEE sources.
Source: Computed from INSEE individual data.

The results obtained confirm our convictions that further gains in life expectancy can only be expected once increasing numbers of people can benefit from the previous gains enjoyed by more robust individuals. If this is the case, in the shift from one generation to the next, an alternative process of homogenization and diversification will occur over time in the population with respect to individual frailty, continually pushing the maximum life span of humankind upwards.

Non-parametric approach

Using individual data from the INSEE file for women aged 98 and over, born in 1870-1879, the survival in days was calculated for each woman. Such detailed information was indispensable in applying the extreme value theory.

Let us recall that the first step in the analysis is to estimate the λ parameter and the relative confidence interval. Thus the hypothesis $\lambda \geq 0$ may be compared with the alternative $\lambda < 0$. Should the first hypothesis be proven, then the maximum life span cannot be estimated. However, should the alternative come to pass, a finite limit to the maximum life span may be assumed.

Figure 4 shows that, for certain k values (number of observations), positive values of λ are included in the confidence interval. Even excluding ages older than (approximately) 102 years (containing wide fluctuations due to the small number of observations), for certain k's, $\lambda \geq 0$ is still included within the confidence interval. Thus the $\lambda \geq 0$ hypothesis may not be ruled out. Focusing on the data used here, no conclusions may be drawn as to whether there is or not a limit to the maximum life span. In statistical terms the analysis does not permit further investigations.

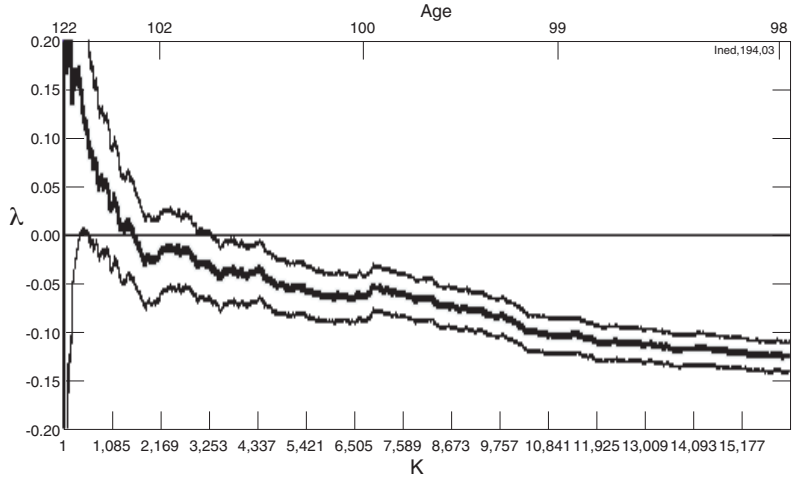

Figure 4.– Estimate (central line) and 95% confidence interval for the extreme value parameter λ (vertical axis) against the number of observations *K* (lower horizontal axis) and the corresponding age (upper horizontal axis)

Source: INSEE individual data, Jeanne Calment included.

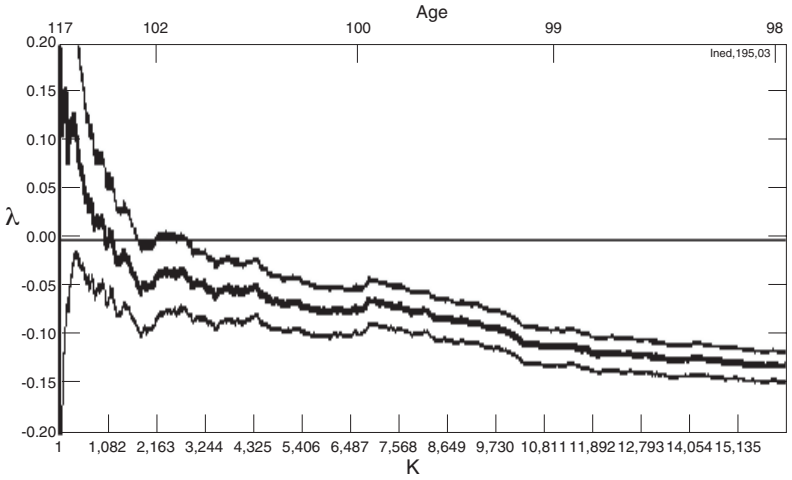

Figure 5.– Estimate (central line) and 95% confidence interval for the extreme value parameter λ (vertical axis) against the number of observations *K* (lower horizontal axis) and the corresponding age (upper horizontal axis)

Source: INSEE individual data, without age at death of Jeanne Calment.

Since Jeanne Calment's age at death is a unique fact — in part still controversial — we repeated the analysis, excluding it from the set. As can be seen from Figure 5, positive values of λ are included in the confidence interval even in this case, and the hypothesis $\lambda \geq 0$ cannot be rejected. As before, we cannot estimate the maximum life span.

Some speculation is nonetheless warranted, particularly in light of the analysis of aggregate data. A tendency in favour of increasing human longevity clearly emerges, without however identifying whether a limit in fact exists. In any case, the analysis performed here focused on changes in the human life span, not on its boundaries. Early results apparently point to the fact that should there be a limit, it is not yet manifest. Nor can a similar conclusion be dismissed in light of the analysis of individual data.

Conclusion

It is more and more obvious that heterogeneity can be an important factor in the dynamics of mortality at the oldest ages. The selection process can explain the observed slowing down in the mortality function at the oldest ages. We tried here to develop several aspects of this analytical framework further.

First, it appeared that it is not only important to try to estimate the effects of heterogeneity in the level of mortality risks at one specific age, but also to look at the heterogeneity in the speed of mortality increase with age, the *slope heterogeneity*. From an application to French cohort data, it appears that slope heterogeneity is probably even more important than level heterogeneity. In particular, the model based on slope heterogeneity may fit the data better and estimate the possible maximum life span more accurately. Of course, these findings are not a description of the reality but are only based on a model which is an oversimplification of much more complex facts, since we made the hypothesis of a population with only two components: frail and not frail. Nevertheless it gives results that fit the observed maximum ages at death very well in the French case.

Second, individual French data allowed us to perform a non-parametric analysis, and the results do not undermine a hypothesized unlimited human life span. Indeed, values of λ greater or equal to 0 that imply that the lifespan might have no limit, are not systematically out of the confidence interval. This not only appears in Figure 4, but also in Figure 5 when the exceptional case of Jeanne Calment is dropped.

Finally, particularly in light of the analysis on aggregate data, a tendency in favour of an increasing human life span clearly emerges, even if it is impossible to provide evidence of the absence of any limit. Should there be one, it is not yet manifest. And the analysis of individual data does not undermine this conclusion.

REFERENCES

AARSSEN Karin, de HAAN Laurens, 1994, "On the maximal life span of humans", *Mathematical Population Studies*, 4(4), pp. 259-281.

ALLARD Michel, LÈBRE Victor, ROBINE Jean-Marie, 1994, *Les 120 ans de Jeanne Calment, doyenne de l'humanité,* Paris, Le Cherche Midi, 162 p.

ALLARD Michel, LÈBRE Victor, ROBINE Jean-Marie, 1998, *Jeanne Calment, from Van Gogh's time to ours, 122 extraordinary years,* New York, W. H. Freeman and Company, 136 p.

BARBI Elisabetta, 1999, *Eterogeneità della popolazione e sopravvivenza umana* : *prospettive metodologiche ed applicazioni alle generazioni italiane 1870-1895,* Florence, Dipartimento Statistico (Ph.D. thesis), 91 p.

BOURGEOIS-PICHAT Jean, 1952, "Essai sur la mortalité biologique de l'homme", *Population*, 7(3), pp. 381-394.

FISHER Ronald, TIPPETT Leonard, 1928, "Limiting forms of the frequency distribution of the largest or smallest member of a sample", *Proceedings of the Cambridge Philosophical Society*, 24, pp. 180-190.

GOMPERTZ Benjamin, 1825, "On the nature of the function expressive of the law of human mortality", *Philosophical Transactions,* Vol. XXVII, pp. 513-585.

GUMBEL Emil J., 1937, *La durée extrême de la vie humaine*, Laval, Barnéoud ; Paris, Hermann, 65 p.

HORIUCHI Shiro, COALE Ansley, 1990, "Age patterns of mortality for older women", *Mathematical Population Studies,* 2(4), pp. 245-267.

KANNISTO Väinö, 1994, *Development of Oldest-Old Mortality, 1950-1990: Evidence from 28 Developed Countries,* Odense (Denmark), Odense University Press (Odense Monographs on Population Ageing, No. 1), 108 p.

KANNISTO Väinö, 1996, *The Advancing Frontier of Survival,* Odense (Denmark), Odense University Press (Odense Monographs on Population Ageing, No. 3), 144 p.

KANNISTO Väinö, LAURITSEN Jens, THATCHER A. Roger, VAUPEL James W., 1994, "Reductions in mortality at advanced ages: several decades of evidence from 27 countries", *Population and Development Review*, 20(4), pp. 793-810.

MANTON Kenneth, STALLARD Eric, TOLLEY Dennis, 1991, "Limits to human life expectancy: Evidence, prospects, and implications", *Population and Development Review*, 17(4), pp. 603-637.

MESLÉ France, VALLIN Jacques, 2002, "Improving the accuracy of life tables for the oldest old: The case of France", *Population*, 57(4-5), pp. 601-630.

THATCHER Roger, 1999, "The long-term pattern of adult mortality and the highest attained age", *Journal Royal Statistical Society (A)*, 162, pp. 5-43.

THATCHER Roger, KANNISTO Väinö, VAUPEL James, 1998, *The Force of Mortality at Ages 80 to 120*, Odense (Denmark), Odense University Press (Odense Monographs on Population Ageing, No. 5), 104 p.

VALLIN Jacques, 1993, *L'avenir de l'espérance de vie*, Paris, INED-PUF (Congrès et colloques, No. 12), 106 p.

VALLIN Jacques, MESLÉ France, 2001a, *Tables de mortalité françaises pour les XIX^e et XX^e siècles et projections pour le XXI^e,* Paris, INED (Données statistiques, No. 4), 102 p. + CD-rom.

VALLIN Jacques, MESLÉ France, 2001b, "Vivre au-delà de 100 ans", *Population et Sociétés*, No. 365, pp. 1-4; an English version, "Living beyond the age of 100" is available on our website www.ined.fr

VAUPEL James W., 1997, "Demographic analysis of ageing and longevity", Paper presented at the *XXIII IUSSP General Population Conference*, Beijing, 17 p.

VAUPEL James W., YASHIN Anatoli I., 1985, "Heterogeneity's ruses: some surprising effects of selection on population dynamics", *American Statistician*, 39, pp. 176-185.

VAUPEL James W., CAREY James R., 1993, "Compositional interpretations of medfly mortality", *Science,* 269, pp. 1666-1667.

VAUPEL James W., JEUNE Bernard, 1995, "The emergence and proliferation of centenarians", in B. Jeune and J.R. Vaupel (eds), *Exceptional Longevity: From Prehistory to the Present*, Odense University Press, pp. 125-169.

VAUPEL James W., MANTON Kenneth, STALLARD Eric, 1979, "The impact of heterogeneity in individual frailty on the dynamics of mortality", *Demography,* 16, pp. 439-454.

VAUPEL James W., CAREY James R., CHRISTENSEN Kaare et al., 1998, "Biodemographic trajectories of longevity", *Science,* 280, pp. 855-860.

VINCENT Paul, 1951, "La mortalité des vieillards", *Population*, 6(2), pp. 181-204.

WILMOTH John, HORIUCHI Shiro, 1999**, "**Do the oldest old grow old more slowly?", *in* Bernard Forette, Claudio Franceschi, Jean-Marie Robine, Michel Allard (eds.)*, The Paradoxes of Longevity,* Berlin, Springer & Verlag, pp. 35-60.

YASHIN Anatoli I., UKRAINTSEVA Svetlana V., DE BENEDICTIS Giovanna et al., 2001, "Have the oldest old adults ever been frail in the past? A hypothesis that explains modern trends in survival", *Journal of Gerontology*, 10(56A), pp. B432-B442.

Elisabetta BARBI, Department of Demography, University of Rome "La Sapienza", e-mail: Barbi@demogr.mpg.de

BARBI Elisabetta, CASELLI Graziella, VALLIN Jacques.– **Trajectories of Extreme Survival in Heterogeneous Populations**

Recent gains in life expectancy among the elderly have noticeably contributed to increasing average life expectancy in developed countries. The old and oldest old are reaching thresholds that were unthinkable 30 or 40 years ago. Are recent gains due to increased longevity among a growing proportion of the population? Or are such gains the harbinger of new frontiers that may announce the further "extension" of the survival curve? Deeper comprehension of underlying mechanisms hinges on models that consider the impact of heterogeneity in individual frailty.

In this paper, we analyse the mortality trajectories of French women born between 1820 and 1879. We applied a classic frailty model and a mixture frailty model accounting for individual differences both in the level of mortality and in the rate of aging. The survival trajectories obtained with these models were used to estimate the maximum life span. Moreover, a non-parametric approach was applied to female centenarians born in France between 1870 and 1879 to estimate the extreme age at death. Results confirm that population heterogeneity can be an important factor in the dynamics of mortality at the oldest ages. In particular, the mixture frailty model may fit the data better and estimate the possible maximum life span more accurately. A tendency towards an increasing human life span clearly emerges, without however establishing whether a limit in fact exists.

BARBI Elisabetta, CASELLI Graziella, VALLIN Jacques.– **Hétérogénéité des générations et âge extrême de la vie**

Les gains récents d'espérance de vie pour les personnes âgées ont considérablement contribué à l'augmentation de l'espérance de vie dans les pays développés. Les personnes âgées et très âgées atteignent des âges qui étaient inimaginables il y a 30 ou 40 ans. Ces progrès d'espérance de vie sont-ils dus à l'augmentation de la longévité d'une part croissante de la population ? Ou bien s'agit-il des signes avant-coureurs de nouveaux sommets qui pourraient annoncer une « extension » de la courbe de survie ? Pour mieux comprendre les mécanismes en jeu, on doit faire appel à des modèles qui prennent en considération l'hétérogénéité de la fragilité individuelle.

Nous analysons ici les trajectoires de mortalité de femmes françaises nées entre 1820 et 1879. Nous appliquons un modèle de fragilité classique et un modèle de fragilité combiné permettant de tenir compte des différences individuelles à la fois pour le niveau de mortalité et pour le rythme de vieillissement. Avec les trajectoires de survie ainsi obtenues, nous avons tenté d'estimer la durée de vie maximale. De plus, une approche non paramétrique a été appliquée aux femmes centenaires nées en France entre 1870 et 1879, afin d'estimer l'âge extrême au décès. Les résultats confirment que l'hétérogénéité de la population est sans doute un facteur important des dynamiques de la mortalité aux très grands âges. En particulier, le modèle de fragilité combiné pourrait permettre une meilleure adéquation avec les données et une plus grande précision de l'estimation de la durée de vie maximale. Nous voyons clairement émerger une tendance à l'augmentation de la durée de vie, sans que l'on puisse cependant établir l'existence ou non d'une limite.

BARBI Elisabetta, CASELLI Graziella, VALLIN Jacques.– **Trayectorias de supervivencia límite en poblaciones heterogéneas**

El aumento de la esperanza de vida de las personas mayores ha contribuido de forma significativa al aumento de la esperanza de vida total en los países desarrollados. Las personas mayores y las muy mayores están alcanzando límites de edad impensables hace 30 o 40 años. ¿Se deben estos aumentos a una mayor longevidad de una proporción creciente de la población o son, por el contrario, el presagio de nuevas fronteras, que indicarían una "extensión" de la curva de supervivencia? Para entender mejor las causas del fenómeno hay que recurrir a modelos que tienen en cuenta el impacto de la heterogeneidad en el nivel de vulnerabilidad de los individuos.

En este artículo analizamos las trayectorias de mortalidad de las mujeres francesas nacidas entre 1820 y 1879. Aplicamos el modelo de vulnerabilidad clásico y una combinación del modelo de vulnerabilidad teniendo en cuenta diferencias individuales del nivel de mortalidad y del ritmo de envejecimiento. Utilizamos las trayectorias de supervivencia que se obtienen con estos modelos para estimar la duración máxima de vida. Aplicamos, además, un modelo no paramétrico a las mujeres centenarias nacidas en Francia entre 1870 y 1879 para estimar la edad límite de defunción. Los resultados confirman que la heterogeneidad poblacional puede ser un factor importante en la dinámica de la mortalidad en edades avanzadas. El modelo de vulnerabilidad combinado se adapta mejor a los datos y permite estimar la duración máxima de vida posible con mayor precisión. La tendencia al aumento de la duración de vida aparece claramente, pero no se puede determinar si existe un límite de supervivencia.