Assessing the rate of ageing
of the human population

Elisabetta Barbi (barbi@demogr.mpg.de)

ASSESSING THE RATE OF AGEING OF THE HUMAN POPULATION

Elisabetta Barbi

January, 2003

1. Introduction

It has been acknowledged that the rate of ageing of the force of mortality slows down at older ages. The age-trajectory of mortality seems to follow a logistic pattern with deceleration at advanced ages (Vaupel et al., 1979; Horiuchi and Coale, 1990; Manton and Vaupel, 1995; Thatcher et al.,1998). The most popular explanation is that individuals from the same cohort differ in their susceptibility to diseases and death. Biological differences as well as life styles, environmental and living conditions may cause individual frailty (Vaupel and Yashin, 1985; Vaupel and Carey, 1993). Accordingly, the observed deceleration in the force of mortality may be the result of selection processes towards low-frailty individuals that is changes in the composition of heterogeneous populations. Thus, the logistic behaviour of the population's force of mortality (observed) may result from individual's forces of mortality (unobserved) which may follow a very different pattern such as an exponential law.

Deeper understanding of the observed mortality trajectory and its underlying mechanisms at individual level hinges on statistical models that take into account the selective survival hypothesis. In 1979, Vaupel, Manton and Stallard devised a *frailty model* for studying mortality that incorporates the concept of individual susceptibility to death in the analysis of survival data. Let

$$\mu(x,z) = z\mu_0(x)$$

be the force of mortality at age x for an individual with unobserved frailty z, where $\mu_0(x)$, the *baseline hazard function*, is the force of mortality at age x for a standard individual, that is an individual with frailty equal to 1. If frailty is gamma distributed with mean 1 and variance σ^2, then the observed force of mortality in a population at age x is:

$$\bar{\mu}(x) = \frac{\mu_0(x)}{\left[1 + \sigma^2 \int\limits_0^x \mu_0(t)dt\right]} = \mu_0(x)\bar{s}(x)^{\sigma^2}$$

where $\bar{s}(x)$ is the observed survivorship function.

The frailty model for univariate data is still widely used in demography because of its simplicity: it exploits observed data at a population level to provide information on what is not observable, that is the degree of heterogeneity in individual frailty. On

the other hand, the univariate frailty model implies some shortcomings and strong assumptions. One of these is the parametric estimation of the age-trajectory of mortality for individuals, that is the parametric specification of the baseline function. It has been recently stressed that there is no real evidence for the shape of the force of mortality for individuals. In the classic frailty model, usually, the baseline hazard function is assumed to follow a Gompertz curve. But alternative assumptions can produce very different results of the degree of the population heterogeneity (Vaupel and Carey, 1993).

The *correlated frailty model* is a possible solution to this not negligible problem when data on twins or other related individuals are available (Yashin et al., 1995). This approach is based on the correlation of the frailties of the related individuals. The information provided by bivariate data permits to elude the assumptions about the parametric form of the age-trajectory of mortality for individuals. Instead, it suggests the non-parametric estimation of this function based on an estimate of the degree of the population heterogeneity, that is the variance of frailty. Applications of this approach to Danish twins showed that the estimated value of the variance in frailty was about 1.5. As a result, the rise of the underlying force of mortality for individuals was even faster than a Gompertz curve (Yashin and Iachine, 1997). It is clear that, given a population mortality curve decelerating at advanced ages, the more rapid the pace of mortality increase with age for individuals, the greater the heterogeneity in frailty. Results on variance in frailty from Italian, French, Swedish, Japanese and US data, obtained by applying the univariate frailty model, where the Gompertz curve was assumed to describe the individual force of mortality, set at lower levels (Manton et al., 1986; Barbi, 1998; Horiuchi and Wilmoth, 1998; Barbi et al., 1999). Further analyses with larger samples of related individuals from different countries are required to shed light on this question. Unfortunately, data at such detailed level are not commonly available. Thus, such a fruitful approach has been only partly exploited.

Another possible solution to avoid the parametric specification of the baseline function, this time suitable for aggregate data, is that devised by Caselli, Vaupel and Yashin in 2000. Here, the value of the variance in frailty is estimated by minimising the variance in rates of mortality improvement across ages and time. In the application to Italian mortality data, from 1965 to 1994 and between ages 50 and 99, the degree of heterogeneity is modelled as a constant for all cohorts involved in the period under study. The estimated value was close to 0,5 for both women and men. This result is apparently inconsistent with those obtained by the parametric approach - where the values are lower both for males and females and the sex pattern is different (women are generally more heterogeneous than men) - as well as with higher values estimated by the correlated frailty model.

The goal of this study is to shed light on the rate of ageing of the human population using easily available data at aggregate level. In the next Section, we describe a simple approach, suitable also for grouped data, to estimate the variance in frailty without assuming any arbitrary functional form of the baseline function. Then we turn to mortality surfaces. We first describe a two-dimensional parametric frailty model and then, generalising the approach illustrated in the next Section, we devise a

two-dimensional non-parametric frailty model. Finally, both the parametric and the non-parametric frailty models are applied to mortality surfaces of Italian extinct or almost extinct cohorts.

2. A non-parametric frailty model for aggregate data

To elude any assumption about the parametric form of the baseline hazard function, additional ancillary information are required. Similarly to the approach of stress experiments (Yashin et. al. 1995), Vaupel et al. (2000) suggest the non parametric estimation of the baseline hazard function using data with proportional-hazard covariates and named the method *covariate identified frailty model*. A reasonable assumption about the relationship between the baseline functions of two different sub-populations permits a non parametric estimation of the baseline function or force of mortality for individuals, based on an estimate of the degree of the population heterogeneity. Although the approach is more feasible with individual data, due to the possibility to exploit a larger number of individual characteristics, there is no reason not to apply such approach to demographic aggregate data.

Following the classic frailty model where frailty is gamma distributed with mean 1 and variance σ^2, let us define

$$\overline{\mu}_A(x) = \mu_{0_A}(x)\overline{s}_A(x)^{\sigma_A^2} \qquad \text{and} \qquad \overline{\mu}_B(x) = \mu_{0_B}(x)\overline{s}_B(x)^{\sigma_B^2}$$

as the forces of mortality of population A and population B, respectively, where, as shown in Section 1, $\mu_{0_A}(x)$ and $\mu_{0_B}(x)$ are the baseline hazard functions, σ_A^2 and σ_B^2 are the variances in frailty and $\overline{s}_A(x)$ and $\overline{s}_B(x)$ are the observed survivorship functions of the two populations.

Now, let us assume that the relationship between the two unobserved forces of mortality for individuals is given by

$$\mu_{0_A}(x) = \alpha\mu_{0_B}(x).$$

In other words, the baseline functions of the populations A and B are assumed to be proportional to each other. Then, it is easy to show that

$$\overline{\mu}_A(x) = \alpha\overline{\mu}_B(x)\overline{s}_B(x)^{-\sigma_B^2}\overline{s}_A(x)^{\sigma_A^2}.$$

Notice that the only parameters of the model to be estimated are the proportionality factor α and the variances in frailty σ_A^2 and σ_B^2.

Comparisons of the hazard functions may be carried out by sex, by cause of death, by geographical origins, by race *etc.*, that is by all those covariates that are commonly recorded at aggregate level. Accordingly, it is theoretically always possible to elude the parametric specification of the force of mortality for individuals even when dealing with grouped data.

3. Frailty models for Mortality Surfaces

Mortality surfaces represent mortality rates as function of age and time. They can well summarise the impact of all those factors - biology, genetics, environment and living conditions – which in fact can differently affect mortality depending on age and time. The analysis of mortality surfaces offers hence the advantage of simultaneously capturing age, period and cohort effects on mortality and their possible hidden interactions more easily than the analysis of age-specific mortality rates at a given time or the analysis of mortality rates over time for a given age (Arthur and Vaupel, 1984; Caselli et al.,1985; Vaupel et al., 1997).

Since 1987, different softwares were devised to perform demographic surfaces (see Vaupel et al., 1997). Mortality surfaces in this paper have been performed using *Lexis*, a program devised by Kirill Andreev (1999).

3.1 Parametric approach

A parametric two-dimensional function, also accounting for selection, for the analysis of mortality surfaces is that devised by Jim Vaupel in 1999.
Let

$$\mu(x,y,z) = z\mu_0(x)exp[-\phi(y)] + c\,exp[-\psi(y)]$$

be the force of mortality at age x and time y for individuals with unobserved frailty z, where $\mu_0(x)$ is the baseline function (the force of mortality for a standard individual), c is a positive constant and the two time-specific functions, ϕ and ψ, denote period effects. If z follows a Gamma distribution with mean 1 at the starting age of observation x_0 and the baseline function is specified by the Gompertz model then, the force of mortality at the population level, at age x and time y, is (Vaupel et al., 1979):

$$\bar\mu(x,y) = a\exp[bx - \phi(y)]\bar{s}(x_0,x,y-x)^{\sigma^2(y-x)} + c\exp[-\psi(y)],$$

where $\bar{s}(x_0,x,y-x)$ is the proportion of the cohort born at time $y-x$ that survived from age x_0 to x and $\sigma^2(y-x)$ is the cohort-specific variance of the Gamma distribution at the starting age of observation x_0. Thus the model is able to decompose mortality surfaces into age, period and cohort patterns.

Italian mortality surfaces have been analysed in this paper via a simplified version of the model where, in order to save degrees of freedom, the two parameters ϕ an ψ are modelled as follows:

$$\phi(x) = \phi y \text{ and } \psi(x) = \psi y,$$

and the heterogeneity in frailty is also assumed to be constant over cohorts. The resulting six-parameter function is still suitable to capture mortality surfaces.

3.2 Non-parametric approach

Exploiting the general result given in Section 2, it is easy now to devise a suitable expression to analyze mortality surfaces in a non-parametric framework. Let us denote

$$\overline{\mu}_A(x,y) = \mu_{0_A}(x,y)\overline{s}_A(x_0,x,y-x)^{\sigma_A^2} \quad \text{and}$$

$$\overline{\mu}_B(x,y) = \mu_{0_B}(x,y)\overline{s}_B(x_0,x,y-x)^{\sigma_B^2}$$

as the surfaces of the force of mortality, defined over age x and time y, of population A and population B respectively, where $\mu_{0_A}(x,y)$ and $\mu_{0_B}(x,y)$ are the unobserved surfaces of force of mortality for individuals and the other parameters have the same meaning as in Section 3.1 but refer to the two populations A and B. If the relationship of the two baseline mortality surfaces is given by:

$$\mu_{0_A}(x,y) = \alpha(y-x)\mu_{0_B}(x,y),$$

where $\alpha(y-x)$ is the proportionality factor relative to the cohort born at time y-x, then the mortality surface of population A is defined by:

$$\overline{\mu}_A(x,y) = \alpha(y-x)\overline{\mu}_B(x,y)\overline{s}_B(x_0,x,y-x)^{-\sigma_B^2}\,\overline{s}_A(x_0,x,y-x)^{\sigma_A^2}.$$

4. Applications to Italian Mortality Surfaces

Age-specific probabilities of dying experienced by the Italian cohorts born between 1873 and 1895, in the range of age from 50 to 98, are depicted in figure 1 for both men and women[1]. The maps offer a panoramic view of the male and female evolution of mortality over cohorts. The shading goes from blue to red as the surfaces rise from low to high mortality.

The period impact of the Second World War is easily traced on the maps by the elongated diagonal that is obviously much more evident for men. As it is clear from the figure, women experienced a more and more fast reduction in mortality at all ages. This trend contrasts with the extremely low decline in male mortality, especially among the older cohorts and at ages above 60-70. As a consequence, women enjoy a considerable advantage over men that, among the youngest cohorts, becomes more and more appreciable even at older ages.

Surfaces of Italian mortality for both men and women have been analyzed, first, via the parametric frailty model and then via the non-parametric approach described respectively in Sections 3.1 and 3.2. The non-parametric approach has been performed through the following comparisons: (1) female mortality *versus* male mortality; (2) male mortality *versus* female mortality; (3) female mortality *versus* total mortality; (4) male mortality *versus* total mortality. Both in the parametric and the non-parametric approaches, mortality functions are estimated using the maximum likelihood method.

---

[1] The data used in this paper include cohort life tables from Caselli's database for ages 50 to 85. After age 85, probabilities of dying have been constructed using Kannisto-Thatcher database on old age mortality (http://www.demogr.mpg.de).

Table 1 shows the results of the fitting process via the parametric approach. As expected, due to the different evolution in mortality between men and women, the period factors have a larger impact on women (ϕ_{female}=0.01734 and ψ_{female}=0.02022; ϕ_{male}=0.00665 and ψ_{male}=-0.00329) than on men. The negative value of the parameter ψ for men is anyway negligible especially when considering the low value estimated for the parameter c. The exponential coefficient of the Gompertz model is higher for women than for men (b_{female}=0.13083 and b_{male}=0.11123). Thus, the force of mortality for the standard individual rises faster for women than for men. As a consequence, as discussed in Section 1, women have a higher degree of heterogeneity in frailty with respect to men (σ^2_{female}=0.14663 and σ^2_{male}=0.09710), in accordance with what has been found in previous studies on different developed countries. However, the estimated values of the variance in frailty are set to quite low levels for both sexes.
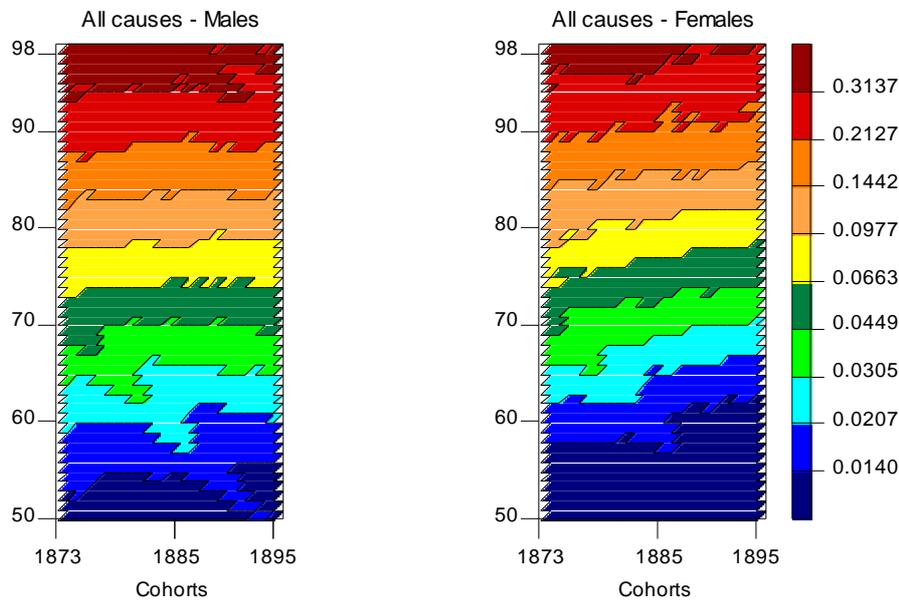
Figure 1: Mortality Surfaces (probabilities of death) for Italian cohorts 1873-1895.

Table 2 shows the results of the fitting process via the non-parametric approach through the four comparisons. In all the applications and for both sexes, the variance in frailty assumes much higher values with respect to the parametric model, determining so a quite rapid pace of the mortality increase with age at individual level. Moreover, the fits with the model where the variances of frailty were imposed to be equal for the two compared populations have led to bad estimates (not shown), assuring us about the statistical significance of their difference. In figure 2 are depicted the trends of the estimated proportionality factors α. They reflect the different mortality evolution over cohorts between men and women discussed above. Thus, a standard woman enjoys

lower mortality risks with respect to a standard man (and, consequently, with respect to the total population) and her advantage rises more and more as cohorts are younger.

Table 1: Parameters with standard errors of the frailty-model for mortality surface fitted to Italian cohorts (1873-1895) from ages 50 to 98.
Parametric approach.

Fitted Parameters	Italy			
	Females		Males	
	est.	s.e.	est.	s.e.
a	0.00417	2.07E-05	0.00567	0.00004
b	0.13083	0.000264	0.11123	0.00033
c	0.00604	4.22E-05	0.00586	0.00005
ϕ	0.01734	6.97E-09	0.00665	0.00000
ψ	0.02022	1.68E-07	-0.00329	0.00000
σ^2	0.14663	0.001586	0.09710	0.00180
Loglikelihood	-21218055		-21259226	

Table 2: Parameters with standard errors of the frailty-model for mortality surface fitted to Italian cohorts (1873-1895) from ages 50 to 98.
Non-parametric approach.

Fitted Parameters	Italian female mortality *vs* Italian male mortality		Italian male mortality *vs* Italian female mortality	
	est.	s.e.	est.	s.e.
$\sigma^2 females$	0.718904	0.007784	0.798088	0.008517
$\sigma^2 males$	0.649594	0.006381	0.720569	0.006958
loglikelihood	-21212313		-21247627	

Fitted Parameters	Italian female mortality *vs* Italian total mortality		Italian male mort. *vs* Italian total mortality	
	Est.	s.e.	est.	s.e.
$\sigma^2 sex$	1.243005	0.023199	0.344210	0.012189
$\sigma^2 total$	1.186763	0.021305	0.358940	0.013660
Loglikelihood	-21207173		-21243717	
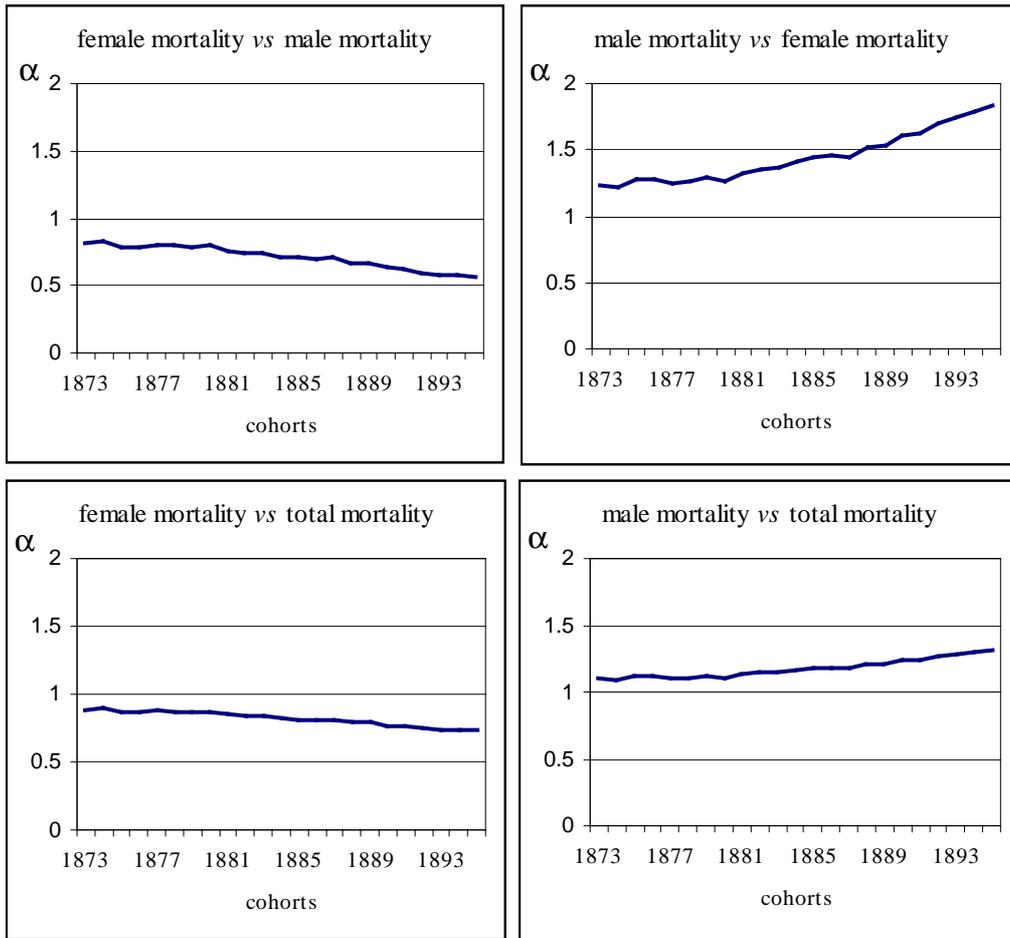
Figure 2: Estimated proportionality factor α of Italian cohorts, 1873-1895.
Non parametric approach.

What is remarkable here, is that values of heterogeneity in frailty seem not only inconsistent with those found via the parametric approach, as expected, but also amongst themselves. While the values of variances for the first two comparisons, being one the opposite of the other, (female mortality *versus* male mortality and male mortality *versus* female mortality) are of the same size order and still comparable (σ^2_{female}=0.718904 and 0.798088; σ^2_{male}=0.649594 and 0.720569), those obtained from the other two applications (female mortality *versus* total mortality and male mortality *versus* total mortality) are enormously different (σ^2_{female}=1.243005; σ^2_{male}=0.344210; σ^2_{total}=1.186763 and 0.358940). The higher incongruence between the two values of the degree of heterogeneity estimated for the total population is probably due to possible non-linear interactions between the variance in frailty for women and that for men.

The log-likelihoods and the Akaike Information Criterion (table 3) show how the better estimates are obtained with the non-parametric approach for both men and women. In this framework, the model comparing the population by sex to the total

population gives the best fit for both sexes and the higher degrees of heterogeneity in frailty for both men and women. Although the high levels reached by the log-likelihoods make this test poorly indicative, we are incline to believe that results from this model are the most reliable for two different reasons.

Table 3: Goodness of fit. Akaike Information Criterion (AIC).

| Models | | AIC | To be compared to |
number of the model	p*	-2loglik. + 2p	model n°
1 – Italian female mortality Parametric approach	6	42436122	3 and 5
2 – Italian male mortality Parametric approach	6	42518464	4 and 6
3 – Female mortality *vs* male mortality Non parametric approach	25	42424676	1 and 5
4 – Male mortality *vs* female mortality Non parametric approach	25	42495304	2 and 6
5 – Female mortality *vs* total mortality Non parametric approach	25	42414396	1 and 3
6 – Male mortality *vs* total mortality Non parametric approach	25	42414396	2 and 4

*p= number of parameters

First, we should consider that the non-parametric approach is very sensitive to the choice of the populations to be compared. The assumption of proportionality between the baseline hazard functions of two populations whose one is part of the other is intuitively the most reasonable. Although, as noted in Section 2, comparisons of the baseline hazard functions may be carried out by a number of different covariates, not all comparisons may make sense. For instance, our application of the method through the comparison Italian mortality *versus* Swedish mortality (not shown) did not give good results. In this perspective, comparisons might then be more reasonable when considering, for instance, regional mortality *versus* mortality at national level or mortality by race *versus* total mortality.

Second, the estimated value of the heterogeneity in frailty, at least for women, match better with that found for the Danish twins (about 1.5) through a completely different method, the correlated frailty model. A similar large value of heterogeneity has been obtained also by Kohlers (2000) by applying the covariate identified frailty model to male and female mortality of Bulgaria. However, the authors consider this value unrealistically high.

Results obtained for Italy seem to confirm that the age-trajectory of the force of mortality for standard individuals may rise even dramatically faster than a Gompertz. The impact of selection would be then already considerable at younger ages and, as a

consequence, the average frailty of the cohorts would be soon reduced. In such a situation, accounting for individual frailty but imposing an arbitrary parametric form for the baseline hazard function may still result in biased demographic life tables. As already noted by Caselli et al. (2000), corrections for unobserved heterogeneity are hence needed not only for the oldest old but also for adult ages.

Acknowledgements

References

Andreev, K. F. (1999), *Demographic surfaces: estimation, assessment and presentation, with application to Danish mortality, 1835-1995*. PhD Dissertation University of Southern Denmark, Odense.

Arthur, W.B., and J.W. Vaupel (1984), Some General Relationships in Population Dynamics, *Population Index* 50(2):214-226.

Barbi Elisabetta, (1998), *Eterogeneità della popolazione e sopravvivenza umana: prospettive metodologiche ed applicazioni alle generazioni italiane 1870-1895*. Ph.D. Dissertation, Firenze.

Barbi Elisabetta, Caselli Graziella, Vallin Jacques (2003), Trajectories of extreme survival in heterogeneous populations, *Population*, Vol 58, n° 1, in press.

Caselli, G., J.W. Vaupel, and A.I. Yashin (1985), Mortality in Italy: Contours of a Century of Evolution, *Genus* 39:1-12.

Caselli Graziella, Vaupel James W., Yashin Anatoli (2000), Longevity, heterogeneity and selection, Paper presented at the XL Riunione Scientifica della Società Italiana di Statistica, Florence, April 26-28.

Horiuchi Shiro, Coale Ansley (1990), Age patterns of mortality for older women, M*athemathical Population Studies,* Vol. 2, n° 4: 245-267.

Horiuchi Shiro, Wilmoth John (1998), Deceleration in the age pattern of mortality at older ages, *Demography,* Vol. 35, n° 4: 391-412.

Kohler, Hans-Peter; Kohler, Iliana (2000), Frailty modelling for adult and old age mortality: the application of a modified DeMoivre Hazard Function to sex differentials in mortality, *Demographic Research* Vol. 3.

Manton Kenneth G., Vaupel James W. (1995), Survival after the age of 80 in the United States, Sweden, France, England and Japan, *New England Journal of Medicine*, Vol. 333: 1232-1235.

Manton Kenneth G., Stallard Eric, Vaupel James W. (1986), Alternative Models for the Heterogeneity of Mortality Risk among the Aged, *JASA*, Vol. 81, n° 395: 635-644.

Thatcher Roger, Kannisto Väinö, Vaupel James (1998), *The force of mortality at ages 80 to 120*. Odense, Monographs on Population Ageing, Vol. 5, Odense University Press, 104 p.

Vaupel James, Carey James R. (1993), Compositional interpretations of medfly mortality, *Science,* Vol. 269: 1666-1667

Vaupel James (1999), The long-term pattern of adult mortality and the highest attained age: discussion on the paper by A.R. Thatcher, *Journal of the Royal Statistical Society/A*, 162:1, 31-43.

Vaupel James, Yashin Anatoli (1985), Heterogeneity's ruses: some surprising effects of selection on population dynamics, *American Statistician*, Vol. 39: 176-185.

Vaupel James, Manton Kenneth, Stallard Eric (1979), The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography,* Vol. 16: 439-454.

Vaupel, James W.; Wang, Zhenglian; Andreev, Kirill F.; Yashin, Anatoli I.(1997), *Population data at a glance*. Odense Monographs on Population, n. 4.

Yashin Anatoli I., Iachine Ivan A. (1997), How frailty models can be used in evaluating longevity limits, *Demography*, n° 1: 31-48.

Yashin Anatoli I., Vaupel James W., Iachine Ivan A. (1995), Correlated Individual Frailty: An Advantageous Approach to Survival Analysis of Bivariate Data, *Mathematical population studies*, Vol. 5, n° 2: 145-159.