Communication proposal for IPC 2021.
Theme: Mortality and Longevity

Mortality above age of 105 using survival analysis in French, Belgian and Quebec populations.

Linh Hoang Khanh Dang[1,2] (*), Carlo Giovanni Camarda[1], France Meslé[1],
Nadine Ouellette[3], Jean-Marie Robine[4,5], Jacques Vallin[1], Michel Poulain[6,7]

[1]French institute for demographic studies
[2]Pole of demography, Université Paris Nanterre
[3]Department of demography, Université de Montréal
[4]French national institute of health and medical research
[5]Ecole pratique des hautes études
[6]IACCHOS, Université catholique de Louvain
[7]EKDK, Tallinn University

(*) Corresponding author: linh.dang@ined.fr

Short abstract.

The biggest challenge when it comes to study mortality at oldest ages is data of reliable quality being rare, data of good quality at individual level are even more sparse. With newly available high quality individual data from France, Belgium and Quebec, we conduct a study on mortality above age 105 using survival analysis method. A preliminary analysis on French dataset shows a continuous increasing trajectory of death risks with a significant male disadvantage even up to these highest ages and an absence of cohort effect. Further exploration with the help of survival analysis's toolbox while extending our research to Belgian and Quebec population will grant us broader view on mortality at oldest ages, where little is known and opinions are often diverged.

Introduction

Studying mortality at extreme old ages has been very challenging, mostly because data of good quality are sparse. Decades of hard work of many research teams offered a new type of data on deaths at oldest ages where validated information at individual level are at disposal, which allows us to make use of methods that could not be of use otherwise. In this paper, we take advantage from high quality individual data from France, Belgium and Quebec and carry out an analysis in proportional hazard setting (Barbi et al., 2018). Our objective is of two-fold: first, to see what this method can unfold on mortality at oldest ages and notably whether it could lend us evidence of a plateau of human mortality for these populations, and second, to dwell more into assessing the gender and cohort effects at these extreme old ages.

Data

For this study, we collect data on birth cohorts whose members were followed from age 105 until extinction from France (1883-1901), Belgium (1891-1904) and Québec (1880-1896). These data are provided at individual level and had gone through a very rigorous process of validation to ensure the accuracy of reported age at death.

Table 1. Distribution of deaths by ages and sexes in each population.

	France		**Belgique**		**Québec**	
Age	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
105	1554	154	147	17	68	10
106	924	78	85	14	41	10
107	505	31	51	0	28	5
108	268	26	26	0	13	2
109	122	10	11	0	7	1
110	58	4	8	1	5	
111	26	1	3	1	2	
112	19		1		2	
113	5				0	
114	3				0	
115	1				1	
Sub total	3485	304	332	33	167	28
Total	3789		365		195	

Consequently, in our analysis framework, there is no occurrence of left-truncated nor right-censored situation, which might be source of biases estimations without proper handling. Right truncation is however present, it is theoretically impossible to be 100% sure that no death at highest ages would not be omitted from our registration. We take into account this element throughout our analysis.

Methods

We model these trajectories from age 105 until death using a parametric proportional hazard model with fixed covariates (gender and birth cohort), in which the baseline hazard function follows the Gompertz model (Gompertz, 1825). The formula for the individual hazard is:

$$h(t_i) = h_0(t_i)\{exp(\beta_1 C_i + \beta_2 M_i)\} = aexp(bt_i)exp(\beta_1 C_i + \beta_2 M_i),$$

where C_i is the individual's birth year minus 1891, while M_i = 1 for males and 0 for females and t_i is the survival duration (in years) of each individual after having attained age of 105. In the coming months we plan to free cohort effects from linearity and test possible separate parameters for different cohorts, possibly grouped.

Given the current setting, β_1 captures a linear cohort effect; β_2 captures the gender effect; a is the initial hazard at starting age 105; and b is the Gompertz slope. The baseline hazard function $h_0(t_i)$ can be then interpreted as the hazard for female subjects born in 1891. With this type of model, a straightforward test can establish the significance of each parameter and evaluate the effect of explanatory variables, e.g., β_2 measures the difference in hazards between males and females, given equal birth cohorts and controlling for age.

Parameters are estimated using the method of maximum of likelihood with right truncation. If ever there are individuals who have survived above age 115 beyond our last date of observation, the right-truncated duration, called R, is the difference between that last date of observation and the date on which these individuals attain age of 105 ($0<t<R$). Given θ the vector of parameters to be estimated, we obtain the log-likelihood function of the following form:

$$logL(t, R; \theta) = log\big(h(t)\big) - H(t) - log\big(1 - S(R)\big)$$

Where h(t) is the hazard function of duration lived, H(t) is the cumulative hazard of duration lived, S(R) is the survival function of right-truncated time.

Regression analysis in our paper is done using R package `flexsurv`, where parameters are estimated using the method of maximum likelihood, with option to take into account the right truncation present in our data scheme.

We then test the hypothesis of having a constant mortality above age 105 against the hypothesis that mortality will keep increasing after 105. In other words, we test the null hypothesis of the Gompertz slope parameter, *b*, being equal to zero against the alternative of a non-zero Gompertz slope parameter. Since the two models are nested, we perform a likelihood ratio test. We also run the same test on all other model specifications to assess the relevance of sex and cohort effects.

The performance of all possible models can be evaluated by the Akaike Information Criteria (AIC). This criterion, which balances a model's fidelity to data and complexity, has the following formula:

$$AIC = -2L + 2k,$$

where L is the maximized likelihood value, given the estimated parameters, and k is the number of parameters being estimated for each model. The absolute values of AIC are not interpretable *per se*, but the difference between AIC scores allows us to rank the associated models' goodness-of-fit while accounting for the number of estimated parameters: the smaller the AIC score, the better the model fits the data. Whereas the AIC provides an overall model assessment, we will also test the statistical significance of the parameters that capture the effects of birth cohort and gender.

Preliminary results

For this extended abstract, we would like to present first our results obtained in French population. Results for Belgian and Quebec population will be provided in the future.

A likelihood ratio test conducted between models under the null hypothesis ($b = 0$) and alternative hypothesis ($b \neq 0$). In full model where both cohort and gender variables are included, a likelihood-ratio test rejects the null hypothesis of a constant hazard after age 105 (p-value = 8.76e-08). A comparison based on AIC also shows that the model with a non-zero Gompertz slope describes the data better (10878.00 vs. 10904.63). Based on reported estimates, Gompertz slope, b, is effectively significant in French data. Parameter estimates for each model specification and its AIC are given in details in Table 2.

Table 2. Parameter estimates in model specifications of different combinations of variables applied to all French individuals born between 1883 and 1901, surviving to age 105, and then followed until death.

Parameter	**Estimate**	**95% CI**	**Log-likelihood**	**AIC**	**Rank**
Model with constant hazard, and cohort and gender effects					
a	0.638	[0.614; 0.664]	-5449.317	10904.63	6
β_1	-0.00017	[-0.00635; 0.006015]			
β_2	0.144	[0.026; 0.261]			
Model with Gompertz hazard, and cohort and gender effects					
a	0.580	[0.550; 0.612]	-5435.002	10878.00	2
b	0.062	[0.040; 0.084]			
β_1	0.00028	[-0.00590; 0.006460]			
β_2	0.156	[0.038; 0.273]			
Model with constant hazard, no covariates					
a	0.645	[0.625; 0.666]	-5452.094	10906.19	7
Model with Gompertz hazard, no covariates					
a	0.589	[0.562; 0.617]	-5438.235	10880.47	3
b	0.061	[0.039; 0.083]			
Model with constant hazard and gender effect					
a	0.638	[0.617; 0.659]	-5449.319	10902.64	5
β_2	0.144	[0.027; 0.261]			
Model with Gompertz hazard and gender effect					
a	0.581	[0.554; 0.609]	-5435.006	10876.01	1
b	0.062	[0.040; 0.084]			
β_2	0.155	[0.038; 0.273]			

Model with constant hazard and cohort effect					
a	0.645	[0.621; 0.670]	-5452.087	10908.17	8
β_1	-0.00035	[-0.00654; 0.00583]			
Model with Gompertz hazard and cohort effect					
a	0.588	[0.559; 0.619]	-5438.234	10882.47	4
b	0.061	[0.039; 0.084]			
β_1	0.00008	[-0.00610; 0.00626]			

We then continue to perform the likelihood-ratio tests between models having different baseline functions without cohort or gender effect (p-value = 1.402e-07), between models with gender effect only (p-value = 8.778e-08), and between models with cohort effect only (p-value = 1.413e-07). We thus find that the Gompertz slope parameter b is positive and statistically different from zero regardless of the model specification, enhancing furthermore our conclusion that there is yet an evidence of mortality plateau above age 105 for French population using this current method.

Besides, according to AIC, we find that models including the Gompertz slope consistently perform better than those assuming constant hazard. Among them, the model with Gompertz baseline hazard and a gender covariate has the lowest AIC, thereby showing that gender still has a significant effect on mortality after age 105. Assuming all else being equal, male subjects have a hazard rate that is 1.168 times higher than their female counterparts (hazard ratio of $exp(0.155) = 1.168$). Hence, according to our data for France, a male disadvantage seems to persist even at the highest ages. No significant cohort effect is detected among any of the model specifications.

Conclusion

As preliminary results, using survival analysis on data of deaths occurred above age 105 in France, we found no evidence for the plateau of human mortality after age 105. Our estimate of Gompertz slope b is statistically different from 0 across all model specifications. This outcome differs from what found on Italian data by Barbi et al. (2018). Unlike in their study, we also found a significant gender effect and no cohort effect on mortality after age 105 in France.

More results from Belgium and Quebec would give us a broader view on this subject, taking the advantage of data availability at individual level, and potentially might give us element to understand actual differences that were observed between these populations. We believe that given many controversies induced by different combinations of data and methods, research done on the base of data combining different countries population would constitute a reasonable direction.

REFERENCES

Aarssen, K. & de Haan, L. (1994). On the maximal life span of humans, *Mathematical Population Studies*, vol.4, n°4, pp.259-281.

Barbi, E., Caselli, G., Vallin,J. (2003). Trajectories of extreme survival in heterogeneous populations, *Population*, vol.58, n°1, pp.43-66.

Barbi, E., Lagona, F., Marsili, M., Vaupel, J.W., Wachter, K.W. (2018). The plateau of human mortality: Demography of longevity pioneers, *Science*, vol.360, n°6396, pp.1459-1461.

Cournil, A., Robine, J-M., Maier, H., Gampe, J., Vaupel, J.W. (2010). The International Database on Longevity: Structure and contents, in Maier H., Gampe J., et al. (eds), *Supercentenarians*, Berlin Heidelberg, Springer (coll. Demographic Research Monographs), pp.31-40.

Dong, X, Milholland, B., Vijg, J. (2016). Evidence for a limit of human lifespan, *Nature*, vol 539, pp.257-259.

Gampe, J., (2010). Human mortality beyond age 110, in Maier H., Gampe J., et al. (eds), *Supercentenarians*, Berlin Heidelberg, Springer (coll. Demographic Research Monographs), pp.231-245.

Gavrilov, L.A. & Gavrilova, N. S. (2011). Mortality measurement at advanced ages: A study of the Social Security Administration Death Master File, *North American Actuarial Journal*, vol.15, n°3, pp. 432 – 447.

Gavrilova, N. S. &Gavrilov, L. A. (2015). Biodemography of old–age mortality in humans and rodents, *Journals of Gerontology: Biological Sciences*, vol.70, n°1, pp. 1- 9.

Gavrilova, N.S, Gavrilov, L.A., Krut'ko, V.N. (2017). Mortality trajectories at exceptionally high ages: A study of supercentenarians, *Living to 100 Monographs* (Living to 100 Symposium, Orlando FL, Jan. 4 -6), Society of Actuaries.

Gbari, S., Poulain, M., Dal, L., Denuit, M. (2017). Extreme value analysis of mortality at the oldest ages: A case study based on individual ages at death, *Living to 100 Monographs* (Living to 100 Symposium, Orlando FL, Jan. 4 -6), Society of Actuaries.

Gompertz, B., (1825), On the nature of the function expressive of the law of the human mortality and on a new mode of determining the value of life contingencies, *Philosophical Transactions of the Royal Statistical Society*, vol.27, n°2, pp.513-583.

Horiuchi, S., Wilmoth, J.R., (1998), Deceleration in the age pattern of mortality at older ages, Demography, vol. 35, n°4, pp.391-412.

Hosmer, JR, D.W. & Lemeshow, S. (1999). Applied survival analysis: regression modeling of time to event, *Wiley series in probability and statistics*, 386p.

Klein, J.P., Moeschberger, M. L. (1997). Survival analysis: Techniques for censored and truncated data, *Springer*, 502p., (coll. Statistics for Biology and Health).

Meslé, F., Vallin, J., Robine, J-M., Desplanques G., Cournil A., (2010). Is it possible to measure life expectancy at 110 in France?, in Maier H., Gampe J., et al. (eds), *Supercentenarians*, Berlin Heidelberg, Springer (coll. Demographic Research Monographs), pp.231-245.

Mills, M. (2011). Introducing survival and event history analysis, *SAGE publications*, 279p.

Ouellette, N., Bourbeau, R., (2014) Measurement of mortality among centenarians in Canada, *2014 Living to 100 Monographs* (Living to 100 Symposium, Orlando FL, Jan. 8 -10), Society of Actuaries.

Ouellette, N. (2016) La forme de la courbe de mortalité des centenaires canadiens-français, Gérontologie et société, vol.38, n°151, pp.41-53.

Ouellette, N., Meslé, F., Vallin, J., Robine, J-M., (forthcoming), Supercentenarians and semi-supercentenarians in France, *in* Maier H., Jeune B., Vaupel J. (eds.). *Exceptional Lifespans*, Springer

Rau, R., Ebeling, M., Peters, F., Bohk-Ewald, C., Missov, T., (2017). Where is the level of mortality plateau? *Living to 100 Monographs* (Living to 100 Symposium, Orlando FL, Jan. 4 -6), Society of Actuaries.

Rootzén, H., Zholud, D. (2017). Human life is unlimitesd – but short, *Extremes*, vol.20, pp.713-728.

Thatcher, R., Kannisto, V., Vaupel, J.W. (1998), *The force of mortality at ages 80 to 120*, Odense, Odense University Press, 104p., (coll. Odense Monographs on Population Aging, vol.5).

Thatcher, R., (1999), The long-term pattern of adult mortality and the highest attained ages, *Journal of the Royal Statistical Society*, vol.162, part 1, pp.5-43.