

## SHORT REPORT

# Determining the lifetime density function using a continuous approach

Rubén Román, Mercè Comas, Lorena Hoffmeister, Xavier Castells

*J Epidemiol Community Health* 2007;**61**:923–925. doi: 10.1136/jech.2006.052639

**Objective:** To apply a continuous hazard function approach to calculate the lifetime density function (LDF) at any age, and to compare the life expectancies derived from the LDF with those obtained with standard life table (SLT) methods.

**Methods:** Age-specific mortality rates were modeled through a continuous hazard function. To construct the cumulative hazard function, appropriate integration limits were considered as continuous random variables. The LDF at any age was defined on the basis of the elemental relationships with the cumulative hazard function. Life expectancies were calculated for a particular set of mortality data using the SLT approach and the expectancy of the LDF defined.

**Applications and comparisons:** The proposed approach was applied using mortality data from the 2001 census of Catalonia (Spain). A Gompertz function was used to model the observed age-specific mortality rates, which fitted the observed data closely. The LDF and the life expectancy, median and standard deviation of the LDF were derived using mathematical software. All differences, in percentages, between the life expectancies obtained from the two methods were 1.1% or less.

**Conclusions:** The LDF gives a wider interpretation of life duration, by extending a deterministic value like life expectancy to a fully informative measure like the LDF.

Life expectancy is a widely used measure in epidemiology that provides a point estimation of the expected remaining lifetime of an individual at a given age.<sup>1–6</sup> It is commonly calculated with standard life table (SLT) methods, usually given for discrete time intervals.<sup>2–4,7</sup> The existing variability in life duration contrasts with life expectancy, which does not truly represent the wide range of possible lifetime values.<sup>3</sup> The lifetime density function (LDF) is a highly informative measure for studying mortality and life duration, but it cannot be derived directly from age-specific mortality rates.<sup>6</sup> A continuous interpretation of lifetime, and consequently of life expectancy, can be performed using a continuous form of the hazard function and treating age conveniently as a continuous variable. The objective of this analysis was to apply a continuous hazard function approach for the calculation of the LDF at any age, and to compare the life expectancies derived from the LDF with those obtained with SLT methods.

## METHODS

### Basis of the method

We will refer to time as a continuous random variable and not as a discrete time interval, in which the future lifetime expresses the amount of time to be lived after a particular age. The probability density function of lifetime given by  $f(x) = \mu(x)S(x)$  is derived on the basis of the relationship between the hazard function  $\mu(x)$  and the survival function given by  $S(x) = \exp(-H(x))$ . These three functions are equivalent in the sense that any two may be derived from the

third.<sup>2,6,8–10</sup> In the expression of the survival function;  $H(x)$  denotes the cumulative hazard function, which is equivalent to the area under the hazard function  $\mu(x)$ . The area under the hazard function was defined by taking the corresponding integration limits ranging from  $x$ , current age of an individual, to  $x + y_x$ , age at death or quantity of time lived from birth to death, where  $X$  and  $Y_x$  are non-negative continuous random variables.

$$H(y_x) = \int_x^{x+Y_x} \mu(\tau) d\tau \quad x, y_x \geq 0 \quad (1)$$

The calculated area will give the risk of dying at a given age  $x$  up to a particular future time  $y_x$ . Substituting the terms in the expression of the probability density function given by  $f(x) = \mu(x)S(x)$  the following expression is obtained:

$$f(y_x) = \exp\left(-\int_x^{x+Y_x} \mu(\tau) d\tau\right) \mu(x+y_x) \quad x, y_x \geq 0 \quad (2)$$

Where  $\mu(\tau)$  denotes a continuous form of the hazard function. The hazard is a rate and thus it is non-negative and has no upper bound. Under some circumstances, the observed mortality rate can be modeled as a parametric continuous hazard function if proper time units and death as the event of interest are considered.

The LDF cannot be defined solely by its mean and variance because it does not follow a symmetric pattern. Solving equation 2 provides important information in the interpretation of life duration, allowing any statistical measure usually obtained from a density function to be calculated.

## Validation

A specific case was developed. The setting was Catalonia (Spain). Population data of the last census of Catalonia in 2001 and the number of deceased in 2001 were obtained from the Institute of Statistics (Institut d'Estadística de Catalunya; IDESCAT). The calculated mortality rates were modeled through a Gompertz function using the Levenberg–Marquardt non-linear iterative least-squares method.<sup>11–15</sup> Separate models were adjusted for men and women. The Gompertz function was the curve that best fitted our observed data.<sup>11–14, 16–20</sup>

To compare SLT calculations and the approach proposed, life expectancies were calculated by applying both methods to the population and mortality data from Catalonia described above. The differences between the life expectancies predicted by the two approaches were calculated.

## APPLICATIONS AND COMPARISONS

Representation of the observed mortality rates and the modeled Gompertz hazard function showed that the estimated function

**Abbreviations:** LDF, lifetime density function; SLT, standard life table

**Table 1** Life expectancy (age at death) from standard life table calculations

Age (years)	Males					Females				
	Life table		Lifetime density function			Life table		Lifetime density function		
	Life expectancy	Life expectancy	Median age at death	Lifetime SD	% Difference in life expectancy	Life expectancy	Life expectancy	Median age at death	Lifetime SD	% Difference in life expectancy
5	77.00	77.63	79.68	12.56	0.82	83.75	83.70	85.41	10.39	0.05
10	77.06	77.64	79.69	12.52	0.76	83.79	83.71	85.41	10.38	0.10
15	77.11	77.67	79.69	12.45	0.72	83.83	83.71	85.41	10.37	0.14
20	77.30	77.71	79.70	12.36	0.53	83.92	83.72	85.41	10.35	0.25
25	77.54	77.77	79.72	12.24	0.30	84.00	83.73	85.41	10.31	0.32
30	77.74	77.86	79.74	12.06	0.15	84.10	83.75	85.42	10.26	0.42
35	78.00	77.99	79.78	11.83	0.01	84.23	83.78	85.43	10.19	0.53
40	78.30	78.19	79.85	11.51	0.13	84.38	83.83	85.44	10.07	0.65
45	78.72	78.49	79.97	11.09	0.30	84.58	83.92	85.47	9.89	0.77
50	79.23	78.91	80.15	10.56	0.41	84.83	84.07	85.51	9.64	0.90
55	79.95	79.51	80.46	9.90	0.56	85.16	84.31	85.60	9.28	1.00
60	80.82	80.34	80.94	9.10	0.60	85.56	84.68	85.76	8.80	1.03
65	82.00	81.48	81.69	8.16	0.64	86.09	85.25	86.04	8.15	0.97
70	83.42	82.98	82.82	7.12	0.52	86.84	86.11	86.54	7.34	0.84
75	85.23	84.91	84.46	6.00	0.37	87.84	87.34	87.39	6.37	0.57
80	87.49	87.32	86.68	4.87	0.19	89.30	89.06	88.77	5.28	0.27
85	90.33	90.23	89.54	3.79	0.11	91.45	91.34	90.83	4.15	0.12
90	93.70	93.60	92.97	2.82	0.10	94.33	94.21	93.65	3.07	0.12
95	97.74	97.41	96.90	2.01	0.34	98.05	97.66	97.17	2.12	0.40

Life expectancy (age at death), median and standard deviation (SD) from the lifetime density function, and percentage difference in life expectancy between standard life table calculations and the continuous lifetime calculation proposed. Data from Catalonia, 2001.

almost overlaps with the observed mortality rates for both men and women (data not shown).

The lifetime probability density functions at any current age are represented in fig 1 and show a great asymmetry, being left skewed and with a flattened slope for ages under 50 years for men and 60 years for women. As age increases over these values, the density becomes more right skewed and leptokurtic.

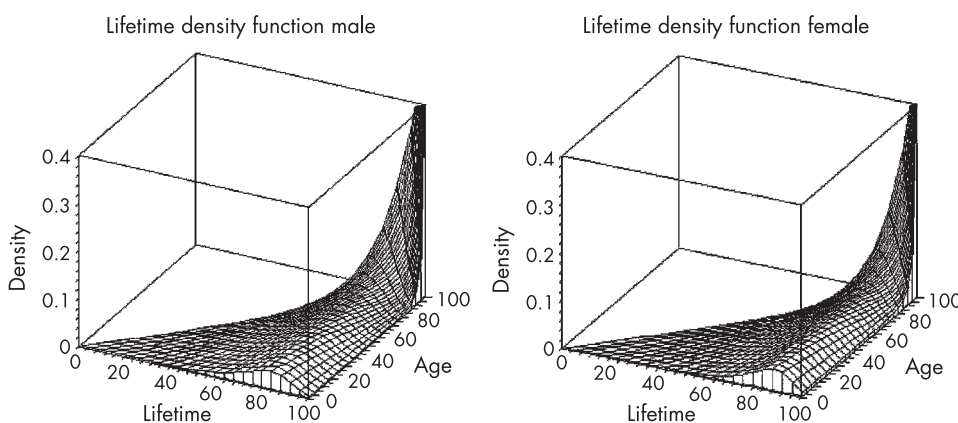
Table 1 compares the life expectancies obtained with the approach proposed with those obtained from the SLT method. Life expectancies were calculated for all age values between one and 100 but, for reasons of space, only ages at five-year intervals are presented in the table; however, the results refer to all the data. All differences were 1.1% or less. For women, the largest difference was found in the middle-age range (approximately 55–65 years), in which a maximum difference of 0.90 (84.51 versus 85.41) more years for the age of 58 years was estimated with the SLT method. This difference diminished gradually as younger and older ages were considered. For males the largest differences were found in the youngest age groups (1–15 years of age) and in those approximately 60 years of age. A maximum difference of 0.70 (77.62 versus 76.92) more years was estimated by the proposed approach for infants aged one

year, and a difference of 0.53 (81.47 versus 82.00) more years was estimated by the life tables for people aged 65 years.

As the present study concerns asymmetric forms of the density function, calculation of the median and standard deviation of the distribution was especially relevant and is included in table 1. Median values were higher than mean values for ages less than 60 years for men and under 75 years for women, reflecting the left skewness of the distribution for these ages. For ages over these values the opposite relationship was found but was less pronounced. Standard deviations were larger in males than in females for ages under 65 years, for ages

**What this paper adds**

The study of mortality is of great interest for epidemiologists, demographers and statisticians. Calculation of the lifetime distribution function at any age using a continuous approach gives useful information for the interpretation of mortality and life duration beyond the usual life expectancy calculated through standard life tables.



**Figure 1** Density function of life duration for men and women. Mortality data from Catalunya 2001.

## Policy implications

The applied procedure, which is based on age-specific mortality rates, gives the opportunity to study life duration more accurately, because all the information about the distribution of lifetime is contained in its density function. Obtaining density functions is essential in several statistical methods such as Bayesian analysis or simulation.

over 65 years women had larger standard deviations. The large standard deviations obtained show the great variability involved with human mortality and life duration.

## DISCUSSION

Life expectancies obtained from the LDF were valid as they were similar to those calculated using SLT methods. The applied method allows the characterisation of the lifetime distribution, which seems essential on account of the level of uncertainty found in life expectancy.<sup>6</sup> The LDF was shown to be non-symmetric and to have different shapes depending on the current age considered. In addition to its utility in the interpretation of mortality, the LDF might be of great interest in situations in which a priori distributions are needed, such as Bayesian statistics or stochastic processes. A poorly estimated hazard function would, however, lead to biased estimations.<sup>2 10 21</sup>

In conclusion, the LDF provides a wider interpretation of life duration by extending a deterministic point estimation such as life expectancy to a totally informative measure like the density function of lifetime values.

## ACKNOWLEDGEMENTS

The authors would like to thank José Miguel Martínez from the Universitat Pompeu Fabra for his support during the writing of this paper.

## Authors' affiliations

Rubén Román, Mercè Comas, Lorena Hoffmeister, Xavier Castells, Evaluation and Clinical Epidemiology Department, Institut Municipal d'Assistència Sanitària (IMAS), Barcelona, Spain

Funding: This work was supported by grants from the Catalan Agency for Health Technology Assessment and Research (CAHTA; 089/07/2000) and the Fondo de Investigación Sanitaria (FIS; PI020365) and the Networks of Excellence (IRYSS G03/202 and RCESP C03/09).

Competing interests: None.

Correspondence to: Xavier Castells, Evaluation and Clinical Epidemiology Department, Institut Municipal d'Assistència Sanitària (IMAS), Passeig Marítim 25–29, 08003, Barcelona, Spain; xcastells@imas.imim.es

Accepted 2 March 2007

## REFERENCES

- 1 Veugelers PJ, Kim AL, Guernsey JR. Inequalities in health. Analytic approaches based on life expectancy and suitable for small area comparisons. *J Epidemiol Commun Health* 2000;**54**:375–80.
- 2 Bongaarts J, Feeney G. Estimating mean lifetime. *Proc Natl Acad Sci U S A* 2003;**100**:13127–33.
- 3 Silcocks PB, Jenner DA, Reza R. Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities. *J Epidemiol Commun Health* 2001;**55**:38–43.
- 4 Miller BG, Hurley JF. Life table methods for quantitative impact assessments in chronic mortality. *J Epidemiol Commun Health* 2003;**57**:200–6.
- 5 Chenet L, McKee M, Otero A, et al. What happened to life expectancy in Spain in the 1980s? *J Epidemiol Commun Health* 1997;**51**:510–14.
- 6 Chiang CL. *The life table and its applications*. Malabar, FL: Krieger, 1984:193–218.
- 7 Chung SJ. Formulas expressing life expectancy, survival probability and death rate in life table at various ages in US adults. *Int J Biomed Comput* 1995;**39**:209–17.
- 8 Chung SJ. Comprehensive life table of computer-assisted predictive mathematical relationship between age and life expectancy, survival probability or death rate in US adults. *Comput Methods Programs Biomed* 1997;**52**:67–73.
- 9 Silverstein MD, Albert DA, Hadler NM, et al. Prognosis in SLE: comparison of Markov model to life table analysis. *J Clin Epidemiol* 1988;**41**:623–33.
- 10 Wilmoth JR. In search of limits. In: Wachter KW, Finch CE, eds. *Between Zeus and the salmon: the biodemography of longevity*. Washington, DC: National Academy Press, 1997:38–64.
- 11 Gompertz B. On the nature of the function expressive of the law of human mortality and on a new mode of determining life contingencies. *Philos Trans R Soc Lond A* 1825;**115**:513–85.
- 12 Haybittle JL. The use of the Gompertz function to relate changes in life expectancy to the standardized mortality ratio. *Int J Epidemiol* 1998;**27**:885–9.
- 13 Neilson S, Robinson I. Reinterpreting mortality statistics: some uses of Gompertzian analysis in epidemiological research. *J Clin Epidemiol* 1993;**46**:1063–9.
- 14 Beck JR, Pauker SG. The markov process in medical prognosis. *Med Decis Making* 1983;**3**:419–58.
- 15 Marquardt D. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J Appl Math* 1963;**11**:431–41.
- 16 Weibull W. A statistical distribution function of wide applicability. *J Appl Mechanics* 1951;**18**:293–7.
- 17 Matsushita S, Hagiwara K, Shiota T, et al. Lifetime data analysis of disease and aging by the Weibull probability distribution. *J Clin Epidemiol* 1992;**45**:1165–75.
- 18 Henderson R, Jones M, Stare J. Accuracy of point predictions in survival analysis. *Stat Med* 2001;**20**:3083–96.
- 19 Flanders WD, Kleinbaum DG. Basic models for disease occurrence in epidemiology. *Int J Epidemiol* 1995;**24**:1–7.
- 20 Reeves GK. The overall distribution of survival times for U.K. AIDS patients. *Philos Trans R Soc Lond B Biol Sci* 1989;**325**:147–51.
- 21 Vaupel JW. Trajectories of mortality at advances ages. In: Wachter KW, Finch CE, eds. *Between Zeus and the salmon: the biodemography of longevity*. Washington, DC: National Academy Press, 1997:17–37.