STATISTICS IN MEDICINE, VOL. 2, 373–380 (1983)

PARAMETERS OF MORTALITY IN HUMAN POPULATIONS WITH WIDELY VARYING LIFE SPANS

WILLIAM SILER

Carraway Methodist Medical Center, Birmingham, Alabama 35234, U.S.A.

SUMMARY

A three-component, competing-risk mortality model, developed for animal survival data, fits human life table data for all ages over a range of mean life spans from 16 to 74 years. The competing risks are a novel exponentially-decreasing hazard, dominant during immaturity; a constant hazard, dominant during adulthood; and an exponentially increasing Gompertzian hazard, dominant during senescence. By fitting the model to a specific life table using non-linear techniques, estimates of the five model parameters and their standard errors obtain; the proportion of deaths expected from each hazard alone may then be calculated. Preliminary analysis of 13 life tables indicates that for human populations under heavy stress, with very short mean life spans of about 20 years, the three hazard components account for roughly equal numbers of deaths; for modern populations, with mean life spans of about 75 years, nearly all deaths are due to the hazard of senescence. Factor analysis of the correlation matrix of parameter values for the 13 populations shows a two-factor structure. One factor involves only the multiplicative constants (initial values) of the three hazards, but not the hazard rates of change; the second factor involves only the parameters of the immaturity hazard and the rate of acceleration of the senescence hazard, but not the constant hazard nor the multiplicative constant (initial value) of the senescence hazard.

KEY WORDS Environmental stress Competing risks Hazard Mortality Survival curves Life tables

INTRODUCTION

Strong interest has been displayed for hundreds of years in quantifying human mortality. However, complete human survival curves (or, equivalently, life tables) have not until now been completely fitted by a mathematical model.

Pearl and Miner,[1] attempting to find a mathematical law to express animal survival curves, abandoned the attempt, saying 'It appears clear that there is no universal law of mortality.' They called for the collection of more animal survival data. There is no lack of life table data for humans. Hutchinson[2] has provided a good historical summary of such data for humans, animals and even grasses. Following the early life expectancy tables of Ulpian in the third century A.D., and the first life table by Gaunt,[3] Halley[4] in 1693 constructed a still-useful life table for the city of Breslau. A particularly useful compilation is that of Preston, Keyfitz and Schoen.[5] Clark[6] compiled a collection of high-stress life tables. Preston[7] dealt with various patterns of mortality.

Attention has also focused on the problem of constructing a mathematical model for mortality. In 1760 Daniel Bernoulli used the notion of competing risks together with Halley's Breslau life table in an attempt to discern the consequences of eliminating smallpox as a cause of mortality, as described by David.[8] Gompertz[9] proposed an exponentially-increasing hazard which is still used.

0277–6715/83/030373–08$01.00

Received 25 October 1982
Revised 22 February 1983

Makeham[10] added a constant hazard to that of Gompertz, again employing the notion of competing risks. After infant and juvenile mortality are spent, Makeham's model fits human data well; it cannot, however, be used to fit an entire life table. The Weibull hazard has considerable current interest for reasons pointed out by David;[8] these include the ability of the Weibull to furnish an increasing, decreasing or constant hazard, depending on the value of its shape parameter. Pinder, Weiner and Smith[11] used the Weibull in a not-too-successful attempt to fit animal life tables. Noble and Hayes[12] used a sum of Gaussians to approximate human mortality data. In addition to models used in an attempt to fit actual life table data, a number of other models have been proposed for various purposes: these include those of Szilard,[13] Sacher and Trucco,[14] Johnson,[15] Curtis,[16] Burch,[17] Forbes, Sprott and Feldstein,[18] Brown and Forbes,[19,20] Takeda,[21] Woodbury[22] and Turner and Pruitt.[23] None of these, however, except possibly that of Takeda,[21] is suitable for fitting complete human life table data. The model of Scala[24] does fit complete life tables for a number of animal species reasonably well.

Siler[25] proposed a three-component competing-risk model for animal mortality. The first component, dominant in the period of immaturity, is an exponentially decreasing hazard; from this hazard alone there is a non-zero survivorship at infinite time. The second component, dominant during adulthood, is simply a constant. The third hazard, dominant during senescence, is an exponentially increasing Gompertzian hazard. All three components are assumed independent and present concurrently throughout the entire life span. Applied to animal survival data, the model fitted the data well. The Weibull hazard was not employed, since, when adjusted to yield a decreasing hazard for the infant and juvenile period, the net survivorship was zero, making it impossible to fit actual survival data; in fact, infant/juvenile mortality is characterized by a non-zero survivorship. Since the model fitted the animal data well, it was checked to see if it also fitted human data and provided any increased insight into human life table data.

I also attempted in the work here reported to find mortality parameter patterns which might be related to environmental stress. Since no data on the environmental stress factors themselves were directly available, I took the average life span as an inverse measure of stress. Few animal data are available for the same species under different environmental stress conditions; Taber and Dasmann[26] provided an outstanding example of such data. As noted above, however, a wealth of such data is available for humans, and the paper will analyse human survival data in the hope that the analysis may prove illuminating for other species as well.

THE MODEL

The model, as given by Siler,[25] has three independent competing risks with hazards $\mu_i(t)$, survivorships $S_i(t)$ and death density functions $f_i(t)$. Then, with non-subscripted quantities representing total effects, the following well-known relationships hold:

$$\mu(t) = \sum_i \mu_i(t)$$

$$S(t) = \prod_i S_i(t) \tag{1}$$

$$f(t) = \sum_i \left(f_i(t) \prod_{j \neq i} S_j(t)\right)$$

The first hazard represents the hazard of immaturity, exponentially decreasing:

$$\mu_1(t) = a_1 \exp(-b_1 t)$$
$$S_1(t) = \exp\{(-a_1/b_1)[1-\exp(-b_1 t)]\}$$
$$f_1(t) = \mathrm{d}S_1(t)/\mathrm{d}t = a_1 \exp(-b_1 t)\exp\{(-a_1/b_1)[1-\exp(-b_1 t)]\}$$

Although these equations seem perfectly straightforward, and are immediately obtained from the well-known Gompertz hazard equations by changing the sign of the exponent, their behaviour is definitely non-standard. Survivorship at infinite time is non-zero, with

$$S_1(\infty) = \exp(-a_1/b_1)$$

Correspondingly, the area of the 'density function' $f_1(t)$ is less than one. Fortunately, when the first hazard is combined with the second and third using equations (1), all straightens out: net survivorship at infinite time is zero, and the area of the net density function is one. The non-zero survivorship from the first hazard represents a biological fact of life; death from the hazards of childhood is not a sure thing.

Relationships for the second and third hazard components are well-known:

$$\mu_2(t) = a_2$$
$$S_2(t) = \exp(-a_2 t)$$
$$f_2(t) = a_2 \exp(-a_2 t)$$
$$\mu_3(t) = a_3 \exp(b_3 t)$$
$$S_3(t) = \exp(a_3/b_3)[1-\exp(b_3 t)]$$
$$f_3(t) = a_3 \exp(b_3 t)\exp(a_3/b_3)[1-\exp(b_3 t)]$$

In combination, then,

$$\mu(t) = a_1 \exp(-b_1 t) + a_2 + a_3 \exp(b_3 t)$$
$$S(t) = S_1(t)S_2(t)S_3(t)$$
$$f(t) = f_1(t)S_2(t)S_3(t) + S_1(t)f_2(t)S_3(t) + S_1(t)S_2(t)f_3(t)$$

PARAMETER ESTIMATION AND DATA FITS

The estimation of competing risks and of risk parameters when individual survival times are known has received considerable attention, as seen in References 27 and 28. When fitting life tables, however, the non-linear regression approach is indicated.[29] I employed simple unweighted least squares with a Gauss–Newton algorithm.[30,31] Asymptotic parameter standard errors and the parameter intercorrelation matrix were obtained by standard linear model theory.

A distinction must be made between two types of parameter correlation matrices. The first, called here the parameter intercorrelation matrix, results from fitting a single data set. Significant intercorrelations imply only a difficulty in distinguishing between effects of the parameters on the error term and reflect a certain kind of uncertainty in the parameter estimates; such intercorrelations do not imply any relation between true values of the parameters for different data sets. I will refer to correlations among true parameter values for different data sets as *correlations*, as distinct from the *intercorrelations* above.

Human life tables were taken from several sources. The Breslau table was taken from Halley.[4] Hutchinson[2] supplied the tables for Sweden for the years 1757–1763, 1816–1840, 1891–1900, 1931–1940 and 1959. Tables for Stone Age Man and West Indian slaves were taken from Clark.[6]

Preston *et al.*[5] gave data for Chile for 1909 and 1920; Taiwan, 1920 and 1930; and the United States, 1964. Average life spans ranged from 15·8 to 74·1 years. Figure 1 displays typical data fits; RMS errors for the 13 life tables averaged $0{\cdot}9 \pm 0{\cdot}3$ per cent and ranged from 0·48 to 1·50 per cent. Although the departures of the data from the model are small and are approximately normally distributed, the errors are in some cases somewhat systematic, as with the Sweden data for 1757–1763, 1816–1840 and 1891–1900. However, the apparent pattern of these departures is not the same from year to year for the same country, nor from country to country. In view of the lack of a pattern in the patterns themselves, I tend to attribute these apparently systematic deviations to the propagation through the years of annual variation in stress on the populations for which the data were collected rather than to an error in the model, although proof of this would require considerable research.

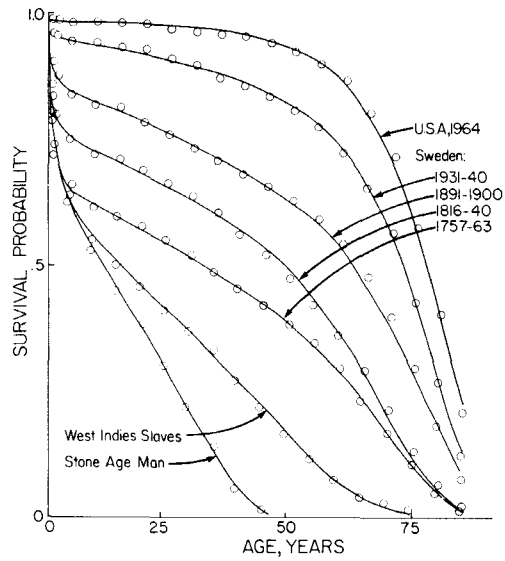

Figure 1. Typical fits of mortality model to human life table data

RESULTS

I now suppose, as did Takeda,[21] that the difference among the survival curves fitted result from environmental stress, which may include cultural factors. To check on the effect of this stress, I first calculated for each life table the proportion of deaths expected from each of the three hazard components. These proportions derive from numerical integration of the equation

$$F_i = \int_0^\infty \mu_i(t) S(t) \mathrm{d}t$$

in which P_i is the proportion of deaths due to the ith hazard, $\mu_i(t)$ is the ith hazard, and $S(t)$ is the survivorship from all hazards. For each life table, I employed the parameter estimates for that table. The results appear in Figure 2. Under the highest stress conditions, i.e. the shortest life span, roughly equal numbers of deaths are due to each of the three hazard components; under the lowest stress, nearly all deaths are attributable to the senescence hazard.

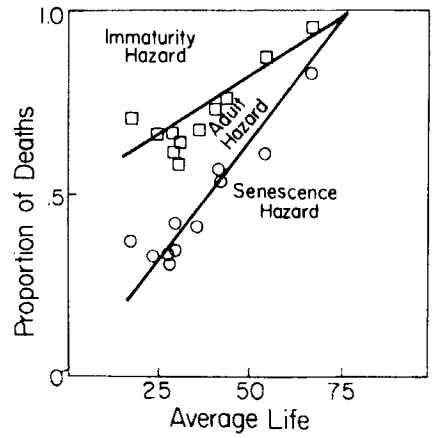

Figure 2. Effect of presumed environmental stress on proportion of deaths from the three hazard components. (The average life span is taken as an inverse measure of environmental stress). The vertical distance from the abscissa to the first (lowest) regression line is the proportion of deaths due to the senescence hazard; the vertical distance between the two regression lines is the proportion of deaths due to the constant hazard, dominant during adulthood; and the vertical distance between the second (higher) regression line and the horizontal line at 1·0 is the proportion of deaths due to the immaturity hazard. Circles are data for the lower regression line; squares are data for the upper regression line.

Table I gives correlations among the mortality components of Figure 2. Table II gives the correlations among the five mortality parameters and the average life span. With only 13 life tables and incompletely screened input data, Table II merits only a tentative interpretation. I used log values as a compensation for non-linear relationships. It is striking that all the parameters correlate at least moderately with average life span except b_3, the rate of acceleration of the senescence hazard. However, b_3 does correlate significantly with both a_1 and b_1, the infant mortality parameters.

I conducted a factor analysis with varimax rotation of the parameter correlation matrix of Table II using the squared multiple correlation coefficient to estimate communalities. The resulting factor loadings appear in Table III. Factor 1 involves only the three initial hazards, or a values; the rates at which these initial hazards are altered are not involved. Factor 2 involves the infant mortality parameters, a_1 and b_1, and the rate of acceleration of the senescence hazard, b_3.

Table I. Correlation coefficients: components of mortality

	Infant	Adult	Senes
Avglife	−0·848	−0·829	0·968
	<0·01	<0·01	<0·01
Infant	1	0·486	−0·812
		<0·1	<0·01
Adult	0·486	1	−0·904
	<0·1		<0·01
Senes	−0·812	−0·904	1
	<0·01	<0·01	

Infant = proportion of deaths due to immaturity hazard, $a_1 \exp(-b_1 t)$
Adult = proportion of deaths due to constant hazard, a_2
Senes = proportion of deaths due to senescence hazard, $a_3 \exp(b_3 t)$
Avglife = average life span.
Probability > $|r|$ shown beneath correlation coefficient.

Table II. Correlation coefficients: parameters of mortality

	Log(a_1)	Log(b_1)	Log(a_2)	Log(a_3)	Log(b_3)
Avglife	−0·67*	0·47	−0·86*	−0·82*	0·07
	<0·01	0·09	<0·01	<0·01	0·80
Log (a_1)	1	0·25	0·39	0·68*	−0·62*
		0·39	0·21	<0·01	0·02
Log (b_1)	0·25	1	−0·18	−0·24	−0·54*
	0·39		0·57	0·41	0·05
Log (a_2)	0·39	−0·18	1	0·39	0·16
	0·21	0·57		0·21	0·63
Log (a_3)	0·68*	−0·24	0·39	1	−0·36
	<0·01	0·41	0·21		0·21
Log (b_3)	−0·62*	−0·54*	0·16	−0·36	1
	0·02	0·05	0·63	0·21	

Total hazard = $a_1 \exp(-b_1 t) + a_2 + a_3 \exp(b_3 t)$
Probability > $|r|$ shown beneath correlation coefficient
Correlation coefficient starred if significant at 0·05 level

Table III. Factor loadings—after varimax rotation

Variable	Factor 1	Factor 2
Log (a_1)	0·75	0·50
Log (b_1)	−0·22	0·72
Log (a_2)	0·56	−0·22
Log (a_3)	0·83	0·08
Log (b_3)	−0·25	0·82

Total hazard = $a_1 \exp(-b_1 t) + a_2 + a_3 \exp(b_3 t)$

DISCUSSION

It seems clear from Figure 1 that the proposed model is capable of fitting a wide variety of human life tables with about 1 per cent RMS error and, therefore, of parameterizing such tables with reasonable parsimony.

Although tentative, the factor structure of the parameter correlation matrix appears reasonable, although one would not expect this structure *a priori*. The first factor involves the three hazard multiplicative constants, but not the rates of change; the second factor involves the rates of change, but not the multiplicative constants, except for the hazard of immaturity. The first hazard, however, is characterized by strong intercorrelation between its rate of change and its multiplicative constant; this artefactual intercorrelation is confounded with any real correlation of these two parameter among the life tables analysed; it remains possible, although unproved, that the second factor involves only the rates of change, but not the multiplicative constants. The effect of increasing stress on a population is apparently such as to increase the magnitude of the hazards in time-independent fashion, and to *slow* the rates of change of the time-varying hazard components. The first factor seems to correspond to the stress imposed on the population, and the second to the

way in which the population responds biologically to this stress. This is speculative; analysis of a much wider data base is clearly in order.

REFERENCES

1. Pearl, R. and Miner, J. R. 'Experimental studies on the duration of life. XIV. The comparative mortality of certain lower organisms', *Quarterly Review of Biology*, **10**, 60–79 (1935).
2. Hutchinson, G. E. *An Introduction to Population Ecology*, Yale University Press, New Haven, Connecticut, 1978.
3. Gaunt, J. *Natural and Political Observations Mentioned in a Following Index and Made Upon the Bills of Mortality, with Reference to the Government, Religion, Trade, Growth, Air, Diseases and the Several Changes of the Said City*, John Martyn, London, 1662.
4. Halley, E. 'An estimate of the degrees of mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslau; with an attempt to ascertain the price of annuities upon lives', *Philosophical Transactions of the Royal Society (London)*, **17**, 596–610 (1693).
5. Preston, S. H., Keyfits, N. and Schoen, R. *Causes of Death: Life Tables for National Populations*, Seminar Press, New York and London, 1972.
6. Clark, C. *Population Growth and Land Use*, Macmillan, London and St. Martins Press, New York, 1968.
7. Preston, H. *Mortality Patterns in National Populations*, Academic Press, New York, 1976.
8. David, H. A. 'Parametric approaches to the theory of competing risks', in Proschan, F. and Serfling, R. J. (eds) *Reliability and Biometry: Statistical Study of Lifelength*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1974.
9. Gompertz, B. 'On the nature of functions expressive of the law of human mortality', *Philosophical Transactions of the Royal Society (London)* **36**, 513–585 (1825).
10. Makeham, W. M. 'On the law of mortality', *Journal of the Institute of Actuaries*, **13**, 325–358 (1860).
11. Pinder, J. E. III, Weiner, J. G. and Smith, M. H. 'The Weibull distributions: a new method of summarizing survivorship data', *Ecology*, **59**, 175–179 (1978).
12. Noble, F. W. and Hayes, J. E. Jr., 'A repetitive analog computer for analysis of sums of distribution functions', *Annals of the New York Academy of Sciences*, **115**, 644–652 (1964).
13. Szilard, L. 'On the nature of the aging process', *Proceedings of the National Academy of Science*, **45**, 30–45 (1959).
14. Sacher, G. A. and Trucco, E. 'The stochastic theory of mortality', *Annals of the New York Academy of Science*, **96**, 985–1007 (1962).
15. Johnson, H. A. Redundancy and biological aging', *Science* **141**, 910–912 (1963).
16. Curtis, H. J. *Biological Mechanisms of Aging*, Charles Thomas, Springfield, Illinois, 1966.
17. Burch, P. R. J. 'Nature's time scale: degenerative disease in man', *Nature*, **216**, 298–299 (1967).
18. Forbes, W. F., Sprott, D. A. and Feldstein, M. 'A model to account for mortality curves of various species', *Journal of Theoretical Biology*, **29**, 293–299 (1970).
19. Brown, K. S. and Forbes, H. F. 'A mathematical model of aging processes', *Journal of Gerontology*, **29**, 46–51 (1974).
20. Brown, K. S. and Forbes, H. F. 'A mathematical model of aging processes', *Journal of Gerontology*, **29**, 401–409 (1974).
21. Takeda, H. 'Stochastic model of aging and death to evaluate the environment', *Medical Journal of Osaka University*, **28**, 97–110 (1977).
22. Woodbury, M. A. and Manton, K. G. 'A random-walk model of human mortality and aging', *Theoretical Population Biology* **11**, 37–48 (1977).
23. Turner, M. D. and Pruitt, K. M. 'A common basis for survival, growth and autocatalysis', *Mathematical Biosciences*, **39**, 113–123 (1978).
24. Scala, C. 'Sulla distribuzione della eta' ala morte dei mammiferi', *Quaderni dell' Istitudodi Statistica*, N **25**, Universita degli Studidi Siena, Siena, Italy, 1976.
25. Siler, W. 'A competing-risk model for animal mortality', *Ecology*, **60**, 750–757 (1979).
26. Taber, R. and Dasmann, R. F. 'The dynamics of three natural populations of the deer, *Odocoileus hemonius columbianus*', in Cox, G. W. (ed), *Readings in Conservation Ecology*, Appleton-Century-Crofts, New York, 1969, pp. 64–69.
27. David, H. A. and Moeschberger, M. L. *The Theory of Competing Risks*, Charles Griffin, London, 1978.
28. Gross, A. J. and Clark, V. A. *Survival Distributions: Reliability Applications in the Biomedical Sciences.* Wiley, New York, 1975.

29. Gehan, E. A. and Siddiqui, M. 'Simple regression methods for survival time studies', *Journal of the American Statistical Association*, **68**, 848–856 (1973).
30. Barr, A. J., Goodnight, J. M., Sall, J. P. and Helwig, J. T. *A User's Guide to SAS 76*, SAS Institute Inc., Raleigh, North Carolina, 1979.
31. Hazelrig, J. B., Turner, M. E. Jr. and Ackerman, E. 'A function minimization package for nonlinear parameter estimation providing readily accessible maximum likelihood estimates', *Computers and Biomedical Research*, **11**, 51–64 (1978).