Twin Research (2000) 3, 51–57

www.nature.com/tr

Genetic nature of individual frailty: comparison of two approaches

Alexander Z Begun[1], Ivan A Iachine[2] and Anatoli I Yashin[3]

[1]Max Planck Institute for Demographic Research, Rostock, Germany
[2]Department of Statistics and Demography, Odense University, Denmark
[3]Duke University, Center for Demographic Studies, Durham, USA

The traditional frailty models used in genetic analysis of bivariate survival data assume that individual frailty (and longevity) is influenced by thousands of genes, and that the contribution of each separate gene is small. This assumption, however, does not have a solid biological basis. It may just happen that one or a small number of genes makes a major contribution to determining the human life span. To answer the questions about the nature of the genetic influence on life span using survival data, models are needed that specify the influence of major genes on individual frailty and longevity. The goal of this paper is to test the nature of genetic influences on individual frailty and longevity using survival data on Danish twins. We use a new bivariate survival model with one major gene influencing life span to analyse survival data on MZ (monozygotic) and DZ (dizygotic) twins. The analysis shows that two radically different classes of model provide an equally good fit to the data. However, the asymptotic behaviour of some conditional statistics is different in models from different classes. Because of the limited sample size of bivariate survival data we cannot draw reliable conclusions about the nature of genetic effects on life span. Additional information about tails of bivariate distribution or risk factors may help to solve this problem. Twin Research (2000) 3, 51–57.

Keywords: frailty models, twins, random hazards

Introduction

The notion of frailty, as a non-observable, hidden susceptibility to death, was introduced in demographic researches at the end of the 1970s. The univariate survival model with proportional hazard and gamma-distributed frailty for the analysis of mortality data was suggested and its properties were studied by Vaupel, Manton and Stallard,[1] Lancaster,[2] Vaupel and Yashin.[3] Hougaard[4] introduced to survival analysis the so-called three-parameter distribution of frailty, which includes as a special case, gamma, inverse-Gaussian, positive stable and degenerate distributions. The first models for bivariate survival data were shared frailty.[5–8] Yashin and Iachine[9] suggested the correlated gamma-frailty model and applied it to genetic analysis of twin data. They showed that at least 50% of frailty variance is conditioned by environmental factors. Different correlated and shared-frailty models were compared on the basis of a general correlated-frailty model with three-parameter distributed frailty.[10] It was shown, that a correlated-frailty model with finite mean of frailty can be regarded as most appropriate to bivariate survival analysis. The choice of frailty distribution for real sample sizes is not very important and, for example, gamma distribution for frailty can be assumed.

However, models with continuously distributed frailty are not the only possible ones. It can be assumed that frailty has a discrete distribution, which may be the case, for example, if hazard is determined by finite number of genes. A two-point model for frailty among sibs based on Mendel's law of inheritance was considered by Mack et al.[11] Vaupel and Tan Qihua[12] estimated the upper limits of the number of longevity genes and found that there are fewer than 400 genetic loci with survival alleles. We suggest here bivariate models with discrete and mixed discrete–continuous frailty distributions. We fit bivariate survival data to Danish twins using these models and then compare our results with previous results for continuously distributed frailty.

Materials and methods

We illustrate our approach using survival data on Danish MZ and DZ twins born in the period 1870–1900 and who both survived until age $x_0 = 30$.

Correspondence: Alexander Z Begun, Max Planck Institute for Demographic Research, Doberaner Strasse 114, D-18057 Rostock, Germany. Tel: (+49) 381 20 81 107; Fax: (+49) 381 20 81 407; E-mail: begun@demogr.mpg.de
Received 30 April 1999; accepted 20 May 1999

https://doi.org/10.1375/twin.3.1.51 Published online by Cambridge University Press

We consider a total of 470 male MZ twin pairs, 475 female MZ pairs, 780 male DZ twin pairs and 835 female DZ twin pairs. Detailed information on these data can be found in Hauge.[13]

We assume that an individual's instantaneous risk of death μ at age t, as measured by the hazard of mortality, depends linearly on frailty Z. Namely, $\mu(t;Z) = Z\mu_0(t)$,where $Z = Z_c + Z_d$, Z_c is the gamma-distributed, continuous, additive part of frailty, and Z_d, which is positive and discretely distributed, is another additive part of frailty, corresponding to finite the number n_g of genotypes; Z_c and Z_d are independent random variables.

For related individuals we assume that life spans T_1 and T_2 are conditionally independent, given frailties Z_1, Z_2, and $Z_{ic} = Y_0 + Y_i$, i = 1,2,

where Y_i are gamma-distributed, independent random variables.

Let $Corr(Z_{1c},Z_{2c}) = \rho$ and $E(Z_{1c}) = E(Z_{2c}) = 1$, $Var(Z_{1c}) = Var(Z_{1c}) = \sigma^2$.

We assume also, that $Z_d = a_{ij} = a_{ji} > 0$ with probabilities $p_{ij} = p_{ji}$, i,j = 1,...,na, $\Sigma_{i,j}\, p_{ij} = 1$. This corresponds to one gene loci with na different possible alleles in two chromosomes. We assume that the first allele in the index pair (i,j) is inherited from the father and allele j is inherited from the mother, both independently. Parents are chosen independently and all persons have the same fertility. In the Hardy-Weinberg equilibrium $p_{ij} = p_i p_j$. If frailty is conditioned by only one beneficial allele, then we have $a = a_0(1-r)^k$ and $a = a_0 - rk$ in the case of k beneficial alleles in the genotype, with correspondingly independent multiplicative and additive action.

Now we can obtain the bivariate survival function in the form $S(x,y) = S_c(x,y)S_d(x,y)$, where

$$S_c(x_1,x_2) = \frac{S_c(x_1)^{1-\rho}\, S_c(x_2)^{1-\rho}}{(S_c(x_1)^{-\sigma 2} + S_c(x_2)^{-\sigma 2} - 1)^{\rho/\sigma 2}}, \quad (1)$$

$$S_c(x) = \{1+H(x)/\sigma^2\}^{-1/\sigma 2}, H(x) = \int_0^x \mu_0(\tau)d\,\tau, \quad (2)$$

$$S_d(x,y) = \sum_{k,l} \tilde{p}_{kl} S_{0d}{}^{\tilde{a}_k}(x) S_{0d}{}^{\tilde{a}_l}(y),\ S_{0d}(.) = e^{-H(.)} \quad (3)$$

and we sum in the third equation all possible combinations of twin genotypes k, l with corresponding weights $\tilde{p}_{kl}$. Denote in the case na = 2

Table 1 Parameter estimates for bivariate survival models with proportional hazards. Equal frailty distributions for males and females. Hardy-Weinberg equilibrium. One beneficial allele with equal for males and females multiplicative action. Danish twins born 1870–1900

Frailty	P	1-r	a22	σ	ρ_{mz}	ρ_{dz}	LogLik	AIC
Discrete sex-linked	0.569 (0.085)	0.414 (0.028)	1.000				–20220.293	40460.586
Discrete autosomal	0.546 (0.082)	0.369 (0.027)	1.000				–20216.913	40453.826
Gamma				1.263 (0.132)	0.494 (0.011)	0.239 (0.073)	–20214.558	40451.116
Mixed 1 genotype		(0.011)	0.010 (0.162)	1.350 (0.015)	0.508 (0.080)	0.237	–20213.801	40451.602
Mixed 3 genotypes	0.013 –	0.708 –	0.009 –	1.355 –	0.506 –	0.235 –	–20213.800	40455.600

P: frequency of beneficial allele, 1-r: coefficient of frailty's decrease.

Table 2 P values of different models

Model	Discrete, sex-linked	Discrete, autosomal	Gamma	Mixed, 1 genotype	Mixed, 3 genotypes
Discrete, sex-linked	0.277[a]				
Discrete, autosomal		0.812[a]			
Gamma			0.649[a]		
Mixed, 1 genotype			0.219	0.916[a]	
Mixed, 3 genotypes		0.183	0.679	0.999	0.966[a]

[a]resulting from a comparison of the general case with the case of equal frailty distributions for males and females and one beneficial allele with equal for males and females multiplicative action for males and females (for the discrete part of frailty)

https://doi.org/10.1375/twin.3.1.51 Published online by Cambridge University Press

genotype (1,1) as 1, genotypes (1,2) and (2,1) as 2, and genotype (2,2) as 3. Let $\tilde{a}_{1,s} = a_{11,s}$, $\tilde{a}_{2s} = a_{12,s} = a_{21,s}$, $\tilde{a}_{3,s} = a_{22,s}$, $\tilde{p}_{1,s} = p_{11,s}$,$\tilde{p}_{2,s} = p_{12,s} + p_{21,s}$,$\tilde{p}_{3,s} = p_{22,s}$, sex s = m,f (males, females). It is not difficult to show that in the special case of autosomal locus, the monozygotic and dizygotic bivariate functions may be expressed as

$$S^{mz}_{d,s}(x,y) = \tilde{p}_{1,m}\tilde{p}_{1,f}S^{\tilde{a}1,s}_{0d,s}(x)S^{\tilde{a}1,s}_{0d,s}(y) +$$
$$0.5\tilde{p}_{1,m}\tilde{p}_{2,f}[S^{\tilde{a}1,s}_{0d,s}(x)S^{\tilde{a}1,s}_{0d,s}(y) + S^{\tilde{a}2,s}_{0d,s}(x)S^{\tilde{a}2,s}_{0d,s}(y)] +$$
$$0.5\tilde{p}_{2,m}\tilde{p}_{1,f}[S^{\tilde{a}1,s}_{0d,s}(x)S^{\tilde{a}1,s}_{0d,s}(y) + S^{\tilde{a}2,s}_{0d,s}(x)S^{\tilde{a}2,s}_{0d,s}(y)] +$$
$$\tilde{p}_{2,m}\tilde{p}_{2,f}[0.5S^{\tilde{a}2,s}_{0d,s}(x)S^{\tilde{a}2,s}_{0d,s}(y) + 0.25S^{\tilde{a}1,s}_{0d,s}(x)S^{\tilde{a}1,s}_{0d,s}(y) +$$
$$0.25S^{\tilde{a}3,s}_{0d,s}(x)S^{\tilde{a}3}_{0d}(y)] + \tilde{p}_{1,m}\tilde{p}_{3,f}S^{\tilde{a}2,s}_{0d,s}(x)S^{\tilde{a}2,s}_{0d,s}(y) +$$
$$\tilde{p}_{3,m}\tilde{p}_{1,f}S^{\tilde{a}2,s}_{0d,s}(x)S^{\tilde{a}2,s}_{0d,s}(y) + 0.5\tilde{p}_{2,m}\tilde{p}_{3,f}[S^{\tilde{a}2,s}_{0d,s}(x)S^{\tilde{a}2,s}_{0d,s}(y) +$$
$$S^{\tilde{a}3,s}_{0d,s}0(x)S^{\tilde{a}3,s}_{0d,s}(y)] + 0.5\tilde{p}_{3,m}\tilde{p}_{2,f}[S^{\tilde{a}2,s}_{0d,s}(x)S^{\tilde{a}2,s}_{0d,s}(y) +$$
$$S^{\tilde{a}3,s}_{0d,s}(x)S^{\tilde{a}3,s}_{0d,s}(y)] + \tilde{p}_{3,m}\tilde{p}_{3,f}S^{\tilde{a}3,s}_{0d,s}(x)S^{\tilde{a}3,s}_{0d,s}(y),$$
$$S^{dz}_{d,s}(x,y) = \tilde{p}_{1,m}\tilde{p}_{1,f}S^{\tilde{a}1,s}_{0d,s}(x)S^{\tilde{a}1,s}_{0d,s}(y) + 0.25(\tilde{p}_{1,m}\tilde{p}_{2,f} +$$
$$\tilde{p}_{2,m}\tilde{p}_{1,f})[S^{\tilde{a}1,s}_{0d,s}(x) + S^{\tilde{a}2,s}_{0d,s}(x)][S^{\tilde{a}1,s}_{0d,s}(y) + S^{\tilde{a}2,s}_{0d,s}(y)] +$$
$$0.25\,\tilde{p}_{2,m}\tilde{p}_{2,f}\,[S^{\tilde{a}2,s}_{0d,s}(x) + 0.5S^{\tilde{a}1,s}_{0d,s}(x) + 0.5S^{\tilde{a}3,s}_{0d,s}(x)]$$
$$[S^{\tilde{a}2,s}_{0d,s}(y) + 0.5S^{\tilde{a}1,s}_{0d,s}(y) + 0.5S^{\tilde{a}3,s}_{0d,s}(y)] +$$
$$0.25(\tilde{p}_{2,m}\tilde{p}_{3,f} + \tilde{p}_{3,m}\tilde{p}_{2,f})[S^{\tilde{a}2,s}_{0d,s}(x) + S^{\tilde{a}3,s}_{0d,s}(x)][S^{\tilde{a}2,s}_{0d,s}(y) +$$
$$S^{\tilde{a}3,s}_{0d,s}(y)] + (\tilde{p}_{1,m}\tilde{p}_{3,f} + \tilde{p}_{3,m}\tilde{p}_{1,f})S^{\tilde{a}2,s}_{0d,s}(x)S^{\tilde{a}2,s}_{0d,s}(y) +$$
$$\tilde{p}_{3,m}\tilde{p}_{3,f}S^{\tilde{a}3,s}_{0d,s}(x)S^{\tilde{a}3,s}_{0d,s}(y).$$

Table 3 Life expectancies and frequencies at age 30 years in the model with beneficial allele 'a'

Genotype	e_{30}		frequency	
	males	females	males	females
AA	31.9	33.5	0.206	0.206
Aa or aA	41.5	43.9	0.496	0.496
aa	50.3	52.9	0.298	0.298
all	42.1	44.4	1.000	1.000

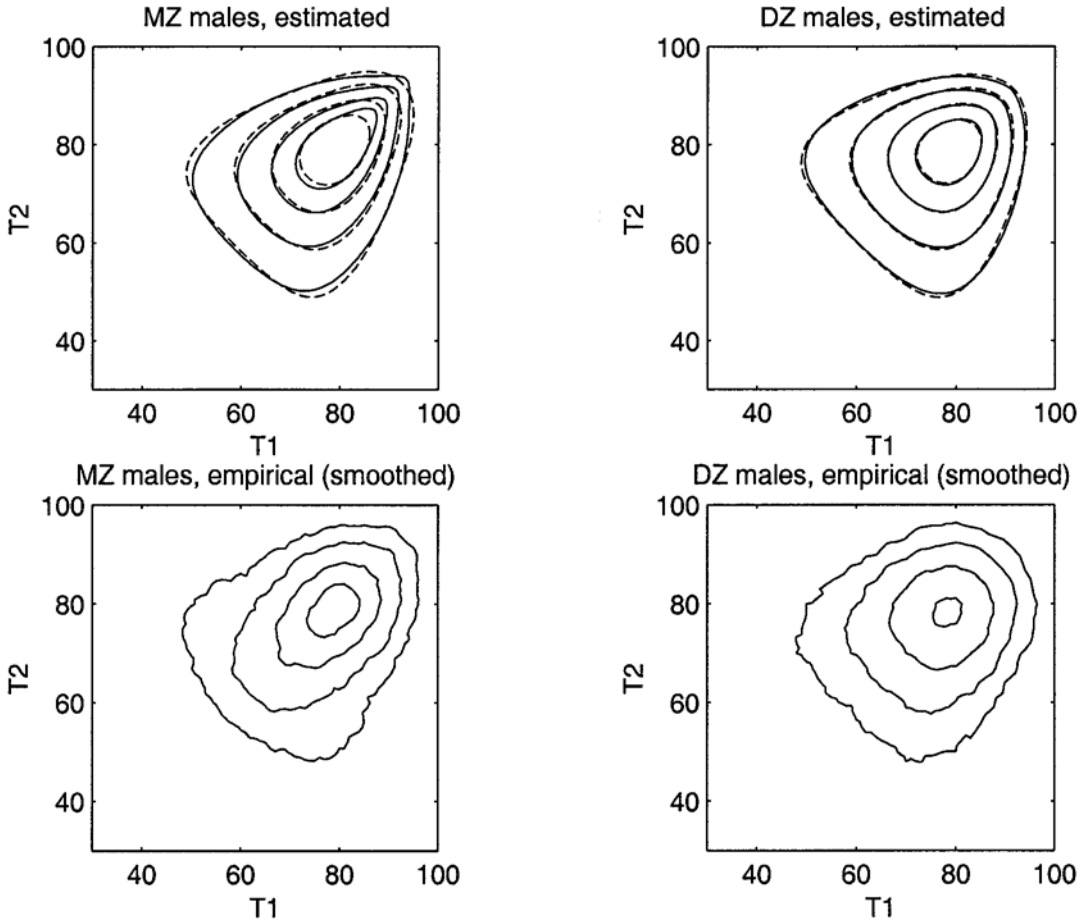

Figure 1 Contour maps of empirical and estimated probability density functions (× 1000) for continuous (solid) and discrete (dashed) frailty models

https://doi.org/10.1375/twin.3.1.51 Published online by Cambridge University Press

Analogous formulae can be written for the sex-linked locus.

Marginal univariate survival functions were approximated with the formula

$$S(x) = \left[1 + s^2\left(ax + \frac{b}{c}\left(e^{cx} - e^{cx_0}\right)\right)\right]^{-\frac{1}{s^2}}.$$

For models with discrete and mixed discrete–continuous frailty we estimated all unknown parameters. These estimates were made for combined MZ and DZ populations for both sexes.

Results

The results of our bivariate analysis are presented in Tables 1–3. All estimates were obtained through maximisation of likelihood function. Unfortunately, standard deviations for the model with mixed discrete–continuous frailty with three genotypes could not be ascertained, because of the non-invertable Hessian. The discrete models with Hardy-Weinberg equilibrium and one beneficial allele which assume multiplicative action of an allele and the discrete models without these last assumptions are nested. That is, we can compare two these types of models by means of the likelihood ratio test. Analysing likelihood values, we conclude that null hypothesis about the Hardy-Weinberg equilibrium and one beneficial allele with multiplicative action which is equal for males and females can be accepted both by autosomal and sex-linked frailty inheritance (Table 2).

We cannot directly compare models with frailty inheritance in autosomal and sex-linked loci since they are not nested. But we can see that the AKAIKE Information Criteria (AIC) is greater in the latter case than in the former (Table 1). In accordance with the likelihood ratio test we accept the null hypothesis about the equality of parameters of gamma-distributed frailties for both males and females (Table 2).

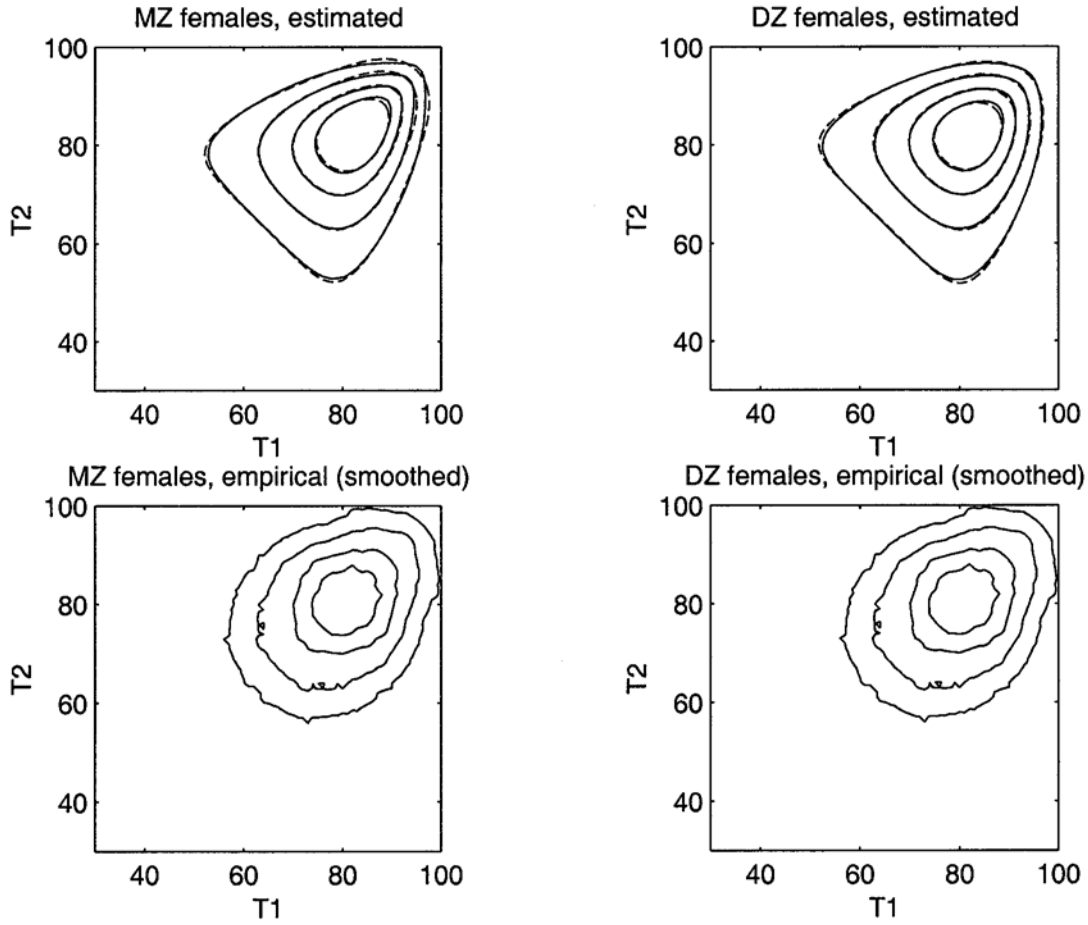

Figure 2 Contour maps of empirical and estimated probability density functions (× 1000) for continuous (solid) and discrete (dashed) frailty models

https://doi.org/10.1375/twin.3.1.51 Published online by Cambridge University Press

We studied two classes of model with mixed frailty: one with one genotype and one with three genotypes (two alleles in one locus). In both cases we can reduce models with assumptions of non-equality of frailty parameters for males and females and non Hardy-Weinberg equilibrium to models with equal frailty parameters for males and females and one beneficial allele with equal, multiplicative action for males and females. In accordance with the likelihood ratio test the mixed frailty model has no advantage over models with purely discrete or purely continuous frailties (Table 2).

We can describe the optimal model with discrete frailty and one beneficial allele in terms of the life expectancies and frequencies of the genotype (Table 3). One can see from this table, that neighbouring genotypes differ from each other by approximately 10 years in life expectancy and that for each genotype female, life expectancy exceeds that of males by about 2 years.

In Figures 1 and 2 one can compare estimated bivariate probability density functions for discrete and gamma-distributed frailties for four populations. These probability density functions were calculated at the age points (x,y), x,y = 30,31,...,100, and then smoothed. It seems that the shapes of probability density functions are very similar in both cases for all populations. In spite of the similar uni- and bivariate fit, the models with discrete and gamma-distributed frailties differ, however, in their underlying hazards, as can be seen in Figure 3. In the model with discretely distributed frailty, underlying hazard does not increase as rapidly as it does in the one with gamma-distributed frailty. One can see in Figure 4 that life-span correlation drastically decreases to zero by the age of 90 in the model with discrete frailty and stabilises at some positive level for the model with continuous frailty.

Discussion

Survival models with continuously distributed frailty are not the only possible ones. As an alternative, one can suggest survival models with discretely distributed frailty. In some sense the latter are more interesting and rich in content. They allow us to measure the influence on mortality of separate genes with specific mechanisms of frailty inheritance, as opposed to the averaged influence of a large number of genes in the case of continuously distributed frailty.

Our analysis reveals a surprising degree of similarity between models with discrete and those with gamma-distributed frailties. This similarity was expressed in the likeness of probability density functions and fits of marginal hazards and maximum likelihood values for all populations we considered. The essential difference between the two models involves the behaviour of underlying hazards and in the asymptotic behaviour or life-span correlations. But these differences are based on the nature of the frailty distribution. In the model with discretely distributed frailty one can see three basic rates of growth of the underlying hazard, corresponding to

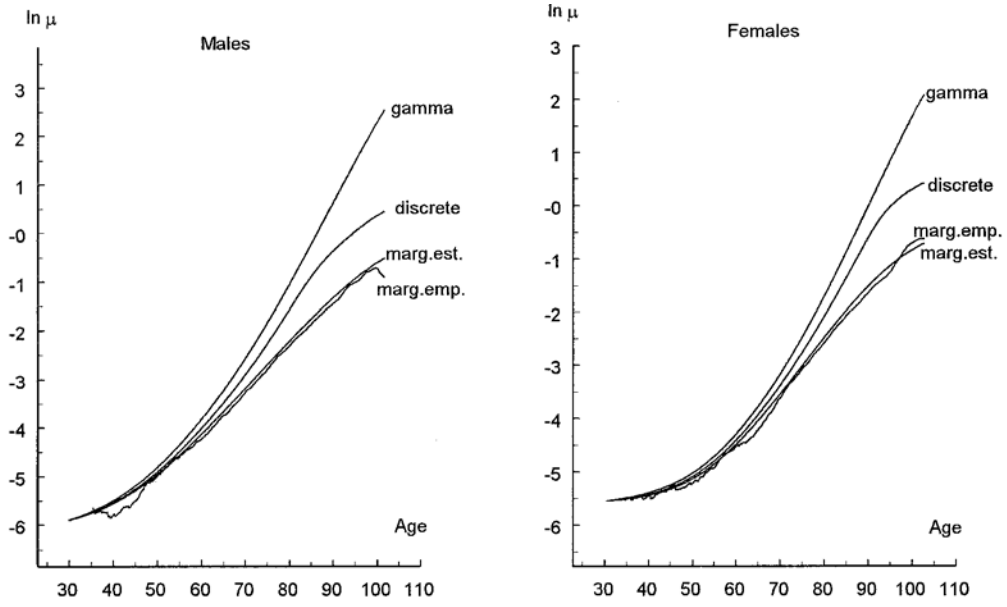

Figure 3 Comparison of the marginal and underlying hazards for gamma and discrete frailty

https://doi.org/10.1375/twin.3.1.51 Published online by Cambridge University Press

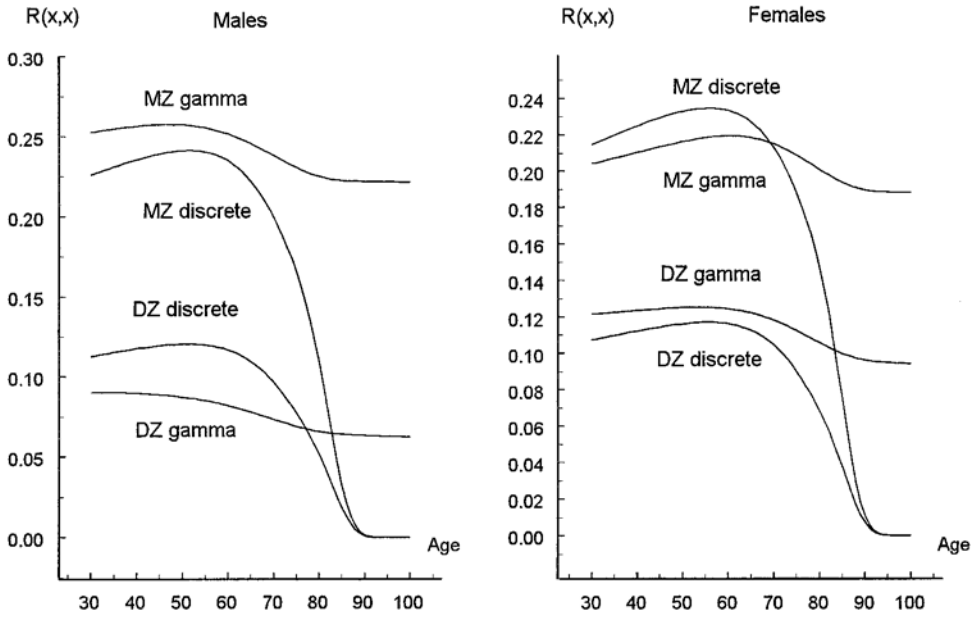

Figure 4 Dynamics of the conditional life span correlations

the three genotypes. In the final period, when the population groups with the two largest values of frailty almost disappear, the underlying hazard is proportional to the marginal hazard, with the coefficient of proportionality equal to the lowest value of frailty.

We observed no deviation from the Hardy-Weinberg equilibrium in a statistical test of this null hypothesis. In addition, we can reduce the model with discrete frailty to the model with the Hardy-Weinberg equilibrium and one beneficial allele with multiplicative action. In our estimation, this beneficial allele is spread with a probability of approximately 1/2 and decreases mortality risk by a factor of about 3 for both sexes. One can see from Table 3 that life expectancy at age 30 years for all genotypes is approximately 2 years longer for females than for males. We can reject the model with frailty inheritance through a sex-linked locus, because it gives greater AIC. This model is probably not as informative as the one with frailty inheritance through an autosomal locus.

Both the model with continuously distributed frailty and the one with discretely distributed frailty are nested in a mixed-frailty model. But in accordance with the likelihood ratio test we cannot reject models with purely discrete and purely continuous frailties. This is probably due to the insufficient size of the sample – in reality we must take into account both the dominant influence of a major gene as well as spread influence of a large number of genes.

Acknowledgements

The authors wish to thank James W Vaupel for useful discussions and Karl Brehmer for help in preparing this paper for publication. This research was partly supported by NIH/NIA grant 7PO1 AG08761-09.

References

1 Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 1979; 16: 439–454.
2 Lancaster T. Econometric methods for duration of unemployment. Econometrica 1979; 47(4): 941–956.
3 Vaupel JW, Yashin AI. Deviant dynamics of death in heterogeneous populations. In: Tuma N (eds). Sociological Methodology, Jossey-Bass: San Francisco, 1985, pp 179–211.
4 Hougaard P. Survival models for heterogeneous populations derived from stable distributions. Biometrika 1986; 73(2): 387–396.
5 Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 1978; 64(1): 141–151.
6 Clayton DG, Cuzick J. Multivariate generalization of the proportional hazards model (with discussion). J Roy Stat Soc, Series A 1985; 148: 82–117.
7 Vaupel JW. Inherited frailty and longevity. Demography 1988; 25: 277–287.
8 Vaupel JW, Yashin AI, Xue L, Holm N, Harvald B. Strategies of modeling genetics in survival analysis. Alternative frailty models applied to data on the life-spans of Danish twins. Population Studies of Aging. Center for Health and Social Policy, Odense University: Denmark, 1993.

https://doi.org/10.1375/twin.3.1.51 Published online by Cambridge University Press

9 Yashin AI, Iachine IA. Environment determines 50% of variability in individual frailty: Results from Danish twin study. Population Studies of Aging. Center for Health and Social Policy, Odense University: Denmark, 1994.
10 Yashin AI, Begun AZ, Iachine IA. Genetic factors in susceptibility to death: a comparative analysis of bivariate survival models. J Epidemiol Biostat 1999; 4(1): 53–60.
11 Mack W, Langholz B, Thomas DC. Survival models for familial aggregation of cancer. Environmental Health Perspectives 1990; 87: 27–35.
12 Vaupel JW, Tan Qihua. How many longevity genes are there? Population Association of America. 1998 Annual Meeting. Chicago, Illinois, 2–4 April 1998.
13 Hauge M. The Danish twin register. In: Mednick SA, Baert AE, Backmann BP (eds). Prospective Longitudinal Research, An Empirical Basis for the Primary Prevention of Psychosocial Disorders. Oxford University Press: London, 1981, pp 218–221.

https://doi.org/10.1375/twin.3.1.51 Published online by Cambridge University Press