STATISTICS IN MEDICINE
Statist. Med. 2003; **22**:3873–3887 (DOI: 10.1002/sim.1669)

The heritability of cause-specific mortality: a correlated gamma-frailty model applied to mortality due to respiratory diseases in Danish twins born 1870–1930‡

Andreas Wienke[1,*,†], Niels V. Holm[2], Kaare Christensen[2], Axel Skytthe[2], James W. Vaupel[1] and Anatoli I. Yashin[1]

[1] *Max Planck Institute for Demographic Research, Rostock, Germany*
[2] *University of Southern Denmark, Odense, Denmark*

SUMMARY

The genetic influence on susceptibility to diseases of the respiratory system and all-cause mortality was studied using data for identical (MZ) and fraternal (DZ) twins. Data from the Danish Twin Register include 1344 MZ and 2411 DZ male twin pairs and 1470 MZ and 2730 DZ female twin pairs born between 1870 and 1930, where both individuals were alive on 1 January 1943. We used the correlated gamma-frailty model. Proportions of variance in frailty attributable to genetic and environmental factors were assessed using the structural equation model approach. For all-cause mortality the correlation coefficients of frailty for MZ twins tend to be higher than for DZ twins. For mortality with respect to respiratory diseases this effect was only seen in females, whereas males showed the opposite effect. Five standard biometric models are fitted to the data to evaluate the magnitude and nature of genetic and environmental factors on mortality. Using the best fitting biometric model heritability for cause of death was found to be 0.58 (0.07) for all-cause mortality (AE-model) and zero for diseases of the respiratory system for males. Heritability was 0.63 (0.11) for all-cause mortality (DE-model) and 0.18 (0.09) for diseases of the respiratory system (DE-model) for females. The analysis confirms the presence of a strong genetic influence on individual frailty associated with all-cause mortality. For respiratory diseases, no genetic influence was found in males and only weak genetic influence in females. The nature of genetic influences on frailty with respect to all-cause mortality is probably additive in males and dominant in females, whereas for frailty with respect to deaths caused by respiratory diseases in females, there are genetic factors present which are caused by dominance. Environmental influences are non-shared with exception of frailty with respect to respiratory diseases in males, where the shared environment plays an important role. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS: Danish twins; correlated gamma-frailty model; respiratory diseases; genetic and environmental factors

*Correspondence to: Andreas Wienke, Max Planck Institute for Demographic Research, Konrad-Zuse-Str. 1, 18057 Rostock, Germany.
†E-mail: wienke@demogr.mpg.de
‡Presented at the Twentieth Annual Conference of the International Society for Clinical Biostatistics, 14–17 September 1999, Heidelberg.

Contract/grant sponsor: NIH/NIA; contract/grant number: 7PO1 AG08761-09

Received September 1999
Accepted September 2003

INTRODUCTION

Recently, there has been an increasing amount of interest placed on assessing the genetic influence on the variation in human life span. Some researchers are interested in specific genes that may influence longevity ('longevity' or 'frailty' genes) [1–3], others are more interested in assessing the overall genetic influence on the variation in human life span. Family studies have shown weak correlations between the age at death of parents and offspring (0.06–0.15) and stronger correlations between the age at death of siblings (0.04–0.35) [4, 5]. The nature of genetic effects responsible for similarities between siblings may be additive or non-additive, whereas factors contributing to parent-offspring resemblance can only be additive. Non-additive genetic effects are not transmitted from parents to offspring because they are the result of gene interaction within (dominance) or between (epistasis) gene loci. Consequently, the higher correlation found among siblings indicates the presence of a non-additive genetic effect, although it may also suggest a higher degree of common environment among siblings.

In twin and adoption studies it is possible to separate the impact of genetic factors and the effect of family environment. Furthermore, questions about the nature of the genetic effect (additive versus non-additive) can be addressed. There are only a few twin studies related to longevity, and most of them showed a higher concordance for age at death in monozygotic twin pairs [6–9], or higher intra-pair differences in life span of DZ twin pairs than in MZ pairs [10]. Only two large studies have been conducted in which a genetic informative cohort has been followed throughout their adult life span. The study by Herskind *et al.* [6], which comprises 2872 Danish twin pairs born between 1870 and 1900, showed that longevity is moderately heritable (0.26 for males and 0.23 for females) and that non-additive genetic factors are important. A Swedish study [7] of 10 505 twin pairs born 1886–1925 concludes that, over the total age range examined, a maximum of around a third of the variance in longevity is attributable to genetic factors and that almost all of the remaining variance is due to non-shared, individually unique environmental factors (note that frailty is the subject of interest in the present study, not longevity as in these two papers just mentioned, which results in markedly different estimates). This analysis was carried out without considering causes of death. But it is well-known that the importance of genetic factors differs for different diseases. Genetic epidemiology seeks to discover the association between genes and diseases. It might be helpful to examine the genetic components of the susceptibility to specific diseases and death rather than longevity. For this purpose we use the correlated gamma-frailty model [11, 12] in this paper, which takes into account the dependence of life spans of related individuals (in our case twins) and allows us to estimate the effect of genetic factors in frailty on a specific cause of death. The approach enables us to combine information about cause of death with data on age at death and to include censored observations. This substantially raises the number of twins (and consequently the statistical power of the analysis) used in our study. For each individual we assume two competing risks of latent times (lifetime related to death due respiratory diseases—lifetime related to death due to all other diseases). Furthermore, we assume that these competing risks are independent. We empirically demonstrate the advantages of the model in the statistical analysis of lifetime data from Danish twins, which were partially used in Herskind *et al*. [6] and McGue *et al*. [10] with a focus on mortality caused by respiratory diseases. The data set was expanded to include 7955 like-sex twin pairs from the birth cohorts 1870–1930, which were followed up from 1 January 1943 until 31 December 1993.

Although respiratory diseases have been largely on the decline in industrialized societies over the last century, it is predicted that they will still play an important role in 2020 [13]. From the point of view of public health, respiratory diseases are greatly underestimated. The early part of the century saw a rapid decline in infectious respiratory diseases [14], mostly from tuberculosis but also from influenza and pneumonia. However, there has also been a steady decline in mortality from most chronic lung diseases. On the other hand, the prevalence of asthma and mortality due to asthma have increased markedly.

MATERIAL AND METHODS

Study population

For our analysis we use survival times of identical (MZ) and fraternal (DZ) male and female twins from the Danish Twin Registry, which was the world's first nation-wide twin registry (established in 1954 by Bent Harvald and Mogens Hauge). This population-based registry includes all twins born in Denmark during the period 1870–1910 and all like-sex pairs born between 1911 and 1930. The birth registers from all 2200 parishes of the relevant calendar years were manually scrutinized to identify all twin births. A search was then carried out for twins, or whenever needed, their closest relatives in regional population registers (in operation since 1924) or other public sources, especially the archives of probate courts and censuses. As soon as a twin was traced, a questionnaire was sent to the twin (if he/she was alive) or to the closest relatives (if not). Questions about phenotypic similarities were included in the questionnaires, so as to assess the zygosity by self-reported similarities. This zygosity classification was compared with laboratory methods (serological markers). The misclassification rate was below 5 per cent [15, 16].

The follow-up procedure traced nearly all twins who did not die or emigrate before the age of 6. For further, detailed information about the construction and the composition of the Danish Twin Registry see Hauge [17].

The registry contains records of 8201 MZ and DZ twin pairs who were born between 1 January 1870 and 31 December 1930 and who were both still alive on 1 January 1943. Consequently, the observations are left truncated. Two hundred and forty six pairs with incomplete cause of death information were excluded, leaving a study population of 7955 pairs. Individuals were followed up to 31 December 1993, and those identified as deceased after that date have been classified here as 'living'. Altogether, we have 1344 male MZ twin pairs and 2411 DZ twin pairs, 1470 female MZ twin pairs and 2730 DZ twin pairs. In addition to the lifetimes, there is information about cause of death for all non-censored observations, i.e. for all included individuals who died before 31 December 1993. For more detailed information about death status, gender, and zygosity of the study population see Table I.

Death status, age at death and cause of death were obtained from the Central Person Register, the Danish Cause-of-Death Register, the Danish Cancer Registry, and other public registries in Denmark. The main source for obtaining information on cause of death was the Death Register at the National Institute of Public Health. Information about cause of death is available from this registry for individuals who died after 1942 [18]. Consequently, cause of death information was included in the twin registry only for twins who died after this year, which results in left truncation in the data set on 1 January 1943.

Table I. Study population by gender, zygosity and status.

	Males				Females			
	MZ		DZ		MZ		DZ	
Both twins dead	731	54.4%	1257	52.1%	622	42.3%	1072	32.4%
One twin alive, co-twin dead	284	21.2%	617	25.6%	332	22.6%	773	28.3%
Both twins alive	329	24.5%	537	22.3%	516	35.1%	885	39.3%
All pairs together	1344	100%	2411	100%	1470	100%	2730	100%

Note: Number of pairs.

Table II. Number of deaths in the study population by gender and zygosity.

	Males				Females			
	MZ		DZ		MZ		DZ	
Deaths from resp. diseases	143	5.3%	203	4.2%	89	3.0%	205	3.8%
Other deaths	1603	59.6%	2928	60.7%	1487	50.6%	2712	49.7%
Alive (censored)	942	35.0%	1691	35.1%	1364	46.4%	2543	46.6%
All individuals together	2688	100%	4822	100%	2940	100%	5460	100%

Note: Number of individuals.

The validity of the twin register was established on the basis of a comparison of information about year of death with the nation-wide Danish Cancer Registry. There was 99 per cent agreement, although both registries were independent [16]. Further data corrections increased this level of agreement to almost 100 per cent.

Mortality

All-cause mortality. After the age of 6, death rates for Danish twins born between 1870 and 1900 are almost the same as those for the same cohorts of the Danish population. The distributions of age at death for MZ twins are close to those of DZ twins for both sexes [19]. This similarity enables us to generalize genetic results from survival models for twins to the whole population with respect to all-cause mortality.

Mortality due to respiratory diseases. Additionally, for our purposes it is necessary to compare the death rates of twins related to diseases of the respiratory system with the respective rates for the general population. The data set contains information whether death was caused by diseases of the respiratory system or not. Respiratory diseases are classified as ICD 470–527 in the sixth and seventh revision and as ICD 460–519 in the eighth ICD revision. The most important subgroups of respiratory diseases are acute lung diseases (acute respiratory infections, influenza, and pneumonia) and chronic lung diseases (bronchitis, emphysema, and asthma). Number of deaths caused by respiratory diseases are given in Table II.

We wished to compare the pattern of death from respiratory diseases for twins with that of the general population. Cause-specific mortality data in five-year age groups (20–24, 25–29, 30–34, . . . , 80–84) are available for the general population going back to 1952 from the WHO

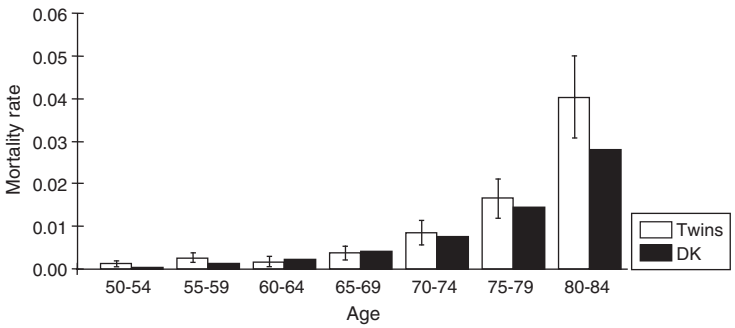

Figure 1. Mortality rates of Danish females with respect to respiratory diseases (Twins = Danish twins, DK = Danish population).

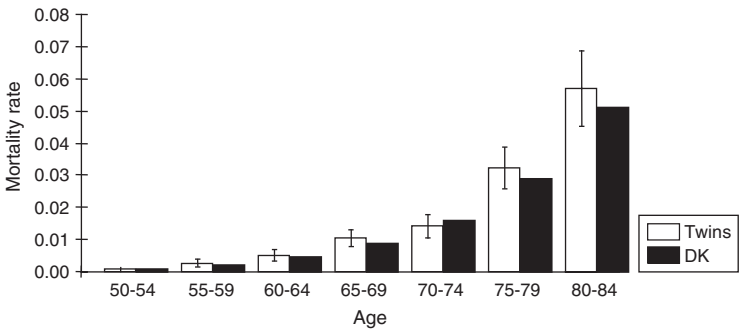

Figure 2. Mortality rates of Danish males with respect to respiratory diseases (Twins = Danish twins, DK = Danish general population).

Mortality Data Base. Death rates for respiratory diseases for males and females, respectively, are shown in Figures 1 and 2. The bars for twin mortality rates include 95 per cent confidence intervals. Confidence intervals were calculated using the binomial formula.

Standardized mortality ratios (SMR) for five-year age intervals were used [20] to detect differences between the mortality pattern of twins and that of the general population. Our analysis shows no significant differences between death rates of female twins and singletons with respect to respiratory diseases (SMR 110 (95 per cent CI: 98–124)). For males a slightly increased mortality of twins compared to the general population was found (SMR 117 (95 per cent CI: 105–130)). These similarities in mortality patterns are fundamental—they allow us to generalize the results from twin studies to the whole population.

STATISTICAL METHODS

Despite the fact that the study population is truncated and partly censored, for a first rough analysis concordance rates, a widely used and accepted indicator of similarities in twins, were computed to facilitate comparisons with other studies (and also with the results of the following analysis combining elements of lifetime analysis and methods of quantitative

genetic). The probandwise concordance rate r was computed as the probability that a twin is affected given that his/her co-twin is affected [21], which means $r = 2C/(2C + D)$, where C and D denote the number of concordant and discordant pairs, respectively. When there is a greater similarity among MZ twin pairs than among DZ twins, we assume an element of heritability. Concordance rates are of limited value in the case of truncated and censored data. For a more powerful approach it is necessary to use the methods of survival analysis. However, univariate lifetime models cannot capture the association between the life spans of related individuals. Consequently, one needs bivariate distributions of dependent lifetimes. The correlated gamma-frailty model [11, 12] (which is an extension of the shared gamma-frailty model [22, 23]) was used to analyse twin lifetimes. It extends the univariate frailty model (for more information about the frailty concept in the univariate case see Vaupel *et al.* [24]). This bivariate lifetime model allows us to include censored observations, which means that we can combine the survival analysis with methods of quantitative genetics.

Let T_i and Z_i $(i = 1, 2)$ be the life times and the frailties of the two individuals of a twin pair. Assume that their individual hazards are represented by the proportional hazards model $\mu(x, Z_i) = Z_i \mu_0(x)$, $i = 1, 2$ with a baseline hazard function $\mu_0(x)$ describing the risk of dying as a function of age. Let the lifetimes T_1 and T_2 be conditionally independent given frailties Z_1 and Z_2, and furthermore a decomposition $Z_1 = Y_0 + Y_1$ and $Z_2 = Y_0 + Y_2$, where Y_0, Y_1, and Y_2 are independent gamma-distributed random variables with $Y_0 \sim G(k_0, \lambda)$, $Y_1 \sim G(k_1, \lambda)$ and $Y_2 \sim G(k_2, \lambda)$. Here k_0, k_1, k_2, λ are non-negative parameters. Obviously, Z_1 and Z_2 are correlated in view of the shared part of frailty Y_0 in both Z_1 and Z_2. To force Z_1 and Z_2 to have the same distribution we assume that shape parameters k_1 and k_2 for the distributions of Y_1 and Y_2 are the same, $k_1 = k_2$. This condition is reasonable for twins, because there is no reason to assume different distributions of frailty in twin partners. Furthermore, we employ the standard assumption that the mean frailty of individuals is one (at the beginning of the follow-up), which means that $EZ_i = 1$, $i = 1, 2$. The common variance is given by $\sigma^2 = 1/\lambda$. Let ρ be the correlation coefficient of Z_1 and Z_2, which is given simply by

$$\rho(Z_1, Z_2) = \frac{k_0}{k_0 + k_1}$$

Because frailties Z_i $(i = 1, 2)$ are usually unobservable, their correlation coefficient used in the methods of quantitative genetics cannot be estimated from the empirical data directly. So a bivariate lifetime model is needed that allows indirect calculation of the parameters. The bivariate gamma distributed frailty with the above mentioned properties was constructed in Yashin *et al.* [11]. The bivariate survival function is given by the following expression:

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

where $S(x)$ is the univariate survival function. Following equation (1) we used a parametric approach by fitting a gamma-Gompertz model to the data, e.g.

$$S(x) = \left(1 + s^2 \frac{\alpha}{\beta}(e^{\beta x} - 1)\right)^{-1/s^2}$$

where α, β, s, correlation coefficient ρ and variance of frailty σ^2 are parameters to be estimated. A comparison with the semiparametric model (where the marginal univariate lifetimes are

non-parametrically estimated using the Kaplan Meier estimator) shows similar results and support this choice of the parametric hazard function (results are not shown here). That is the reason why we use the parametric survival model only for the present analysis. Furthermore, it allows us to use the standard errors calculated in the maximum-likelihood procedure.

As mentioned above, the twin pair data set is not a random selection of twin pairs from the total twin population. Since both members of a twin pair had to be alive on 1 January 1943, the survival times in the data set are sampled from specific conditional distributions. If a twin pair was born in year y (where $y = \{1870, \ldots, 1930\}$), the condition of survival of both twins until the year 1943 implies that both twins had to survive until the age of 1943-y in order to be included in the sample. If the survival times are denoted by X_1 and X_2 with survival function $S(x_1, x_2)$, then the conditional survival function for a twin pair born in year y is:

$$S_y(x_1, x_2) = P(X_1 > x_1, X_2 > x_2 \mid X_1 > 1943\text{-}y, X_2 > 1943\text{-}y) = \frac{S(x_1, x_2)}{S(1943\text{-}y, 1943\text{-}y)} \tag{2}$$

For a combined analysis of MZ and DZ twins, we include two correlation coefficients, ρ_{MZ} and ρ_{DZ}, respectively. These correlations between MZ and DZ twins provide information about genetic and environmental influences on frailty within individuals. Despite the minor differences in the lifetimes of MZ and DZ twins, all other parameters are assumed to be equal for MZ and DZ twins in order to keep the number of parameters in the model as small as possible.

According to standard biometric practice the model assumes no epistasis (genetic interlocus interaction), no gene–environment interaction, and random mating. We assume that the variance of frailty can be expressed as a linear function of unobserved genetic and environmental components: $Z = A + D + C + E$, where A is the additive genetic effect, D is the effect due to dominance (i.e. intralocus interactions), C is the influence of shared environment and E is the effect due to non-shared environment. Resemblance in MZ twins is caused by three factors: additive genetic factors, dominant genetic factors, and shared environmental factors. Any dissimilarities between MZ twins must arise from non-shared environmental factors. In contrast, dissimilarity between DZ twin pairs, who share on average half of the additive and a quarter of the dominant factors, can also be the result of additive and dominant genetic factors.

We calculated the extent to which variance is attributable to genetic and environmental factors using the structural equation model approach, which has been described in detail elsewhere [25].

Five biometric models based on different assumptions about genetic structure were fit to the data: AE, DE, ACE, ADE and CE. In this notation an ACE model refers to the decomposition of frailty $Z = A + C + E$, and a DE model refers to the decomposition $Z = D + E$. ADE, AE, CE models are defined similarly. We use the small letters a^2, d^2, c^2, e^2 to refer to the respective proportions of variance. Standard assumptions about quantitative genetics yield the following relations:

$$\rho_{MZ} = a^2 + d^2 + c^2, \quad \rho_{DZ} = 0.5a^2 + 0.25d^2 + c^2, \quad 1 = a^2 + d^2 + c^2 + e^2 \tag{3}$$

Table III. Number of concordant/discordant pairs and concordance rates.

	Mortality due to respiratory diseases		All-cause mortality	
	MZ	DZ	MZ	DZ
Males				
Both twins died	7	15	731	1257
One twin died	129	173	284	617
Both twins censored	1208	2223	329	537
Concordance rate	0.098	0.148	0.837	0.803
Females				
Both twins died	5	3	622	1072
One twin died	79	199	332	773
Both twins censored	1386	2528	516	885
Concordance rate	0.112	0.029	0.789	0.735

Because of these three relations, only models with three unknown parameters are identifiable. More complex models require mortality information on additional groups of relatives such as parents or children or twins reared apart.

To combine the approach of quantitative genetics with the methods of survival analysis we used the extended correlated gamma-frailty model with genetic and environmental components of frailty. In this approach the genetic and environmental parameters of frailty decomposition are estimated directly by the maximum likelihood method. For more detailed information about this point see Yashin and Iachine [26]. We performed the analysis with statistical software packages SPSS [27] and GAUSS [28].

RESULTS

The number of concordant and discordant twin pairs with respect to death caused by respiratory diseases are given in Table III. Using the classic twin study design we calculate probandwise concordance rates of MZ and DZ twins for a preliminary analysis. For all-cause mortality (both sexes) and for deaths caused by respiratory diseases in females, there were a greater proportion of concordant pairs among the monozygotic twins than among the dizygotic twins, which suggests that there is a genetic component. However, for respiratory diseases in males the opposite effect is observed.

All causes

It is easy to see from Table IV, that for males the ACE and the ADE model converge to the AE model. The estimates of c^2 and d^2 are zero or close to zero, with large standard errors. The AE model is better than the ACE and ADE models according to the likelihood-ratio test. Since not all of our models are nested, we use the Akaike information criterion (AIC) to compare them. According to the AIC the AE model gives the best fit for males, with heritability estimate of 0.58 (0.07). For females, the DE models fits the data best with heritability estimate of 0.63 (0.11). The purely environmental CE model does not fit the data well.

Table IV. Estimates of the components of variance in frailty to all-cause mortality.

Correlated gamma-frailty model (gamma-Gompertz parameterization)						
	σ	a^2	c^2	d^2	e^2	AIC
Males						
ACE	1.333	0.467	0.106		0.427	41304.5334
	(0.178)	(0.118)	(0.091)		(0.071)	
AE	1.364	0.584			0.416	41303.9603
	(0.179)	(0.069)			(0.069)	
ADE	1.364	0.584		0.000	0.416	41305.9603
	(0.179)	(0.069)		(—)	(0.069)	
DE	1.443			0.579	0.421	41324.3124
	(0.201)			(0.068)	(0.068)	
CE	1.277		0.449		0.551	41319.2056
	(0.179)		(0.061)		(0.061)	
Females						
ACE	1.334	0.526	0.000		0.474	15216.7560
	(0.269)	(0.107)	(—)		(0.107)	
AE	1.385	0.508			0.492	15214.7560
	(0.282)	(0.103)			(0.103)	
ADE	1.320	0.428		0.117	0.455	15216.0000
	(0.261)	(0.196)		(0.199)	(0.109)	
DE	1.195			0.626	0.374	15214.2520
	(0.220)			(0.115)	(0.115)	
CE	1.396		0.342		0.658	15219.2080
	(0.370)		(0.091)		(0.091)	

Note: σ^2—Variance of frailty. a^2—Additive genetic factors, c^2—Common environment, d^2—Genetic effects due to dominance, e^2—Individual environment.

Respiratory diseases

The situation for respiratory diseases (Table V) is an entirely different one. For males in models that include genetic components (ACE, ADE, AE, and DE), the standard errors are large compared with the heritability estimators (for example, 0.58 (0.48) in the AE model), indicating that there is no or only small genetic influence in frailty to respiratory diseases. The purely environmental model CE is preferred by the AIC and also by the likelihood ratio test. For females the DE model gives the best fit to the data, with a heritability estimate of 0.18 (0.09). Consistent with these findings are higher concordance rates among male DZ twins than among MZ twins and an inverse result for females (Table III).

Because of the small number of pairs where both twins died of respiratory diseases we performed a combined analysis of both males and females. Despite the fact that the standard deviation σ of frailty differs substantially (3.05 and 9.80 for males and females, respectively) the likelihood ratio test did not indicate significant differences ($\chi^2 = 2.70$, $p = 0.10$). Consequently, assuming σ to be equal for males and females, we tested the hypothesis of equal genetic parameters for males and females in the ACE and ADE model (Table VI). The likelihood ratio test did not detect any evidence of significant differences in the genetic parameters of males and females (ACE model: $\chi^2 = 2.90$, $p = 0.23$; ADE model: $\chi^2 = 2.58$, $p = 0.28$).

Table V. Estimates of the components of variance in frailty to mortality due to respiratory diseases.

	Correlated gamma-frailty model (gamma-Gompertz parameterization)					
	σ	a^2	c^2	d^2	e^2	AIC
Males						
ACE	3.051	0.000	0.422		0.578	4447.5111
	(1.559)	(—)	(0.251)		(0.251)	
AE	2.966	0.581			0.419	4447.6890
	(2.084)	(0.484)			(0.484)	
ADE	2.966	0.581		0.000	0.419	4449.6890
	(2.085)	(0.485)		(—)	(0.485)	
DE	3.500			0.505	0.495	4452.3002
	(2.746)			(0.424)	(0.424)	
CE	3.051		0.422		0.578	4445.5111
	(1.559)		(0.251)		(0.251)	
Females						
ACE	9.499	0.142	0.000		0.858	4056.6840
	(3.112)	(0.072)	(—)		(0.072)	
AE	9.508	0.142			0.858	4054.6840
	(3.113)	(0.072)			(0.072)	
ADE	9.794	0.000		0.185	0.815	4055.3064
	(3.291)	(—)		(0.089)	(0.089)	
DE	9.795			0.185	0.815	4053.3064
	(3.288)			(0.089)	(0.089)	
CE	9.379		0.086		0.914	4057.0444
	(3.092)		(0.047)		(0.047)	

Note: σ^2—Variance of frailty. a^2—Additive genetic factors, c^2—Common environment, d^2—Genetic effects due to dominance, e^2—Individual environment.

Both models imply an AE model as the best fitting model. In the ACE model, the shared environment was zero.

DISCUSSION

This report evaluates data on twin pairs from the Danish Twin Registry. The subjects were followed from 1943 to 1993. Our analysis of mortality from all causes and from diseases of the respiratory system assesses the importance of genetic and environmental factors, and it allows us to calculate heritability estimates and to evaluate the nature of the genetic influence on cause-specific mortality. Our analysis of frailty (liability to mortality from respiratory diseases and from all causes) presented here is based on a relatively large twin population. However, the methodology described in this paper is not restricted to twin analysis. It can be extended to analyse the lifetimes of different relatives, but in this case multivariate correlation's and different family sizes occur and relations (3) have to be change [25].

The major findings of this study are the profound influence of genetic factors on all-cause mortality as well as the moderate influence of genetic factors on diseases of the respiratory system in females. No genetic component was found in frailty to mortality from respiratory

Table VI. Estimates of the components of variance in frailty to mortality due to respiratory diseases. Combined analysis of males and females.

Correlated gamma-frailty model (gamma-Gompertz parameterization)								
	σ	a^2	c^2	d^2	e^2	Log-L	χ^2-statistic	p-value
ACE model								
Males*	4.406	0.000	0.294		0.706			
	(—)	(—)	(—)		(—)			
Females*	4.406	0.222	0.000		0.778	−4241.6140		
	(—)	(—)	(—)		(—)			
Males†	5.450	0.265	0.004		0.731			
	(1.906)	(0.315)	(0.230)		(0.119)			
Females†	5.450	0.265	0.004		0.731	−4243.0618	$\chi^2 = 2.90$	$p = 0.23$
	(1.906)	(0.315)	(0.230)		(0.119)			
ADE model								
Males*	4.877	0.357		0.000	0.643			
	(—)	(—)		(—)	(—)			
Females*	4.877	0.000		0.281	0.719	−4241.7731		
	(—)	(—)		(—)	(—)			
Males‡	5.408	0.272		0.000	0.728			
	(—)	(—)		(—)	(—)			
Females‡	5.408	0.272		0.000	0.728	−4243.0618	$\chi^2 = 2.58$	$p = 0.28$
	(—)	(—)		(—)	(—)			

Note: σ^2—Variance of frailty. a^2—Additive genetic factors, c^2—Common environment, d^2—Genetic effects due to dominance, e^2—Individual environment.

*$\sigma_{\text{males}} = \sigma_{\text{females}}$.

†$\sigma_{\text{males}} = \sigma_{\text{females}}$, $a^2_{\text{males}} = a^2_{\text{females}}$, $c^2_{\text{males}} = c^2_{\text{females}}$.

‡$\sigma_{\text{males}} = \sigma_{\text{females}}$, $a^2_{\text{males}} = a^2_{\text{females}}$, $d^2_{\text{males}} = d^2_{\text{females}}$.

diseases in males. However, a combined analysis of males and females detects a genetic component in frailty to mortality from respiratory diseases.

The classic concordance analysis shows a greater proportion of concordant pairs among the MZ twins than among the DZ twins for all-cause mortality and mortality due to respiratory disorders (the latter for females only). The concordance rates for all-cause mortality are very high because of the long follow-up period.

Estimates of heritability based on different biometric models range from zero (respiratory diseases in males) up to 0.63 (all-cause mortality in females). These estimates are robust in the sense that values derived from the additive AE and non-additive DE model are nearly identical.

The results with respect to all-cause mortality (heritability estimates of 0.58 (0.07) and 0.63 (0.11) for males and females, respectively, from the best-fitting AE and DE model, respectively) confirm the estimates obtained in previous studies: the heritability in frailty is around 0.5 [12, 26, 29] for both sexes, which supports the hypothesis of additive nature of responsible genetic factors [26, 29]. Our estimates of heritability are substantially higher than those found in a previous analysis [6] of mortality in Danish twins born 1870–1900 (approximately 38 per cent of the present sample), which had values between 0.20 and 0.26.

A Swedish study [7] resulted in heritability estimates of around one third. The higher degree of heritability found in our study was expected, because the trait under study was frailty and not life span as in the previous studies. Heritability estimates are always higher for frailty than for life span. Because of the assumption of conditional independence of life spans given their frailties, all genetic and environmental factors responsible for similarities in twin pairs are included in frailty. Apart from this, frailty contains only parts of non-shared environmental factors, other parts being included in the baseline hazard function.

Jarvik *et al.* [10] found in a study of 853 twin pairs (with at least one partner surviving to age 60) mean intra-pair differences in lifespan to be higher in DZ twin pairs (6 years) than in MZ twin pairs (3 years) suggesting a genetic influence on lifetime. A study of 960 adult Danish adoptees [30] supports the hypothesis of genetic influences on premature death; this is consistent with the results in the present study.

While all twins combined have a mortality rate similar to that of the general population, MZ twins show slightly better survival than DZ twins. This difference may be due to selective factors [31].

Our analysis of male and female mortality with respect to respiratory diseases shows no uniform pattern. For males the purely environmental CE models gave the best fit with respect to the AIC, which indicates that mortality due to these diseases have no—or only a weak—genetic component. For females, the CE model gave the worst fit to the data. According to AIC, the DE model is the best, with a heritability estimate of 0.18 (0.09). This is in agreement with the higher concordance rates for male DZ twins compared with MZ twins and the opposite relation for females. A combined analysis of males and females favours an AE model with a heritability estimate of 0.27 (0.31). Perhaps the number of deaths is too small (and standard errors thus too large) to detect the possible gender differences found in the separate analysis.

The finding for males contradicts the conclusion in the literature, that there is a substantially genetic influence in the aetiology of asthma. Asthma is an important cause of death in the group of respiratory diseases. Heritability estimates for asthma obtained by twin studies using questionnaires range from 0.6 to 0.75 [32–34]. For an overview in this area see Koppelman [35]. In the largest twin study [36] to date (undertaken in Finland), heritability of asthma was found to be 0.68 for females and zero for males. We found a similar pattern in the present study. The design of the Finnish study differed from that of previous twin studies of asthma in that the Finnish twins were not contacted directly. Rather, nation-wide health records were computer linked to determine whether asthma was noted on death certificates, hospital discharge records and/or medication-reimbursement records.

However, all these studies are based on asthma rather than on respiratory diseases in general and morbidity was analyzed instead of mortality. Maybe this is one reason for the difference compared to our study.

To our knowledge, the present analysis is the first twin study in respiratory disease mortality in general. Furthermore, our analysis allows us to combine information on cause of death with lifetime information. This survival analysis approach promises to be a more powerful tool than the common analysis of binary traits (ill/healthy or death/alive).

Frailty is a latent variable similar to the conventional approach, which deals with the latent variable liability (usually assumed to be normally distributed) as susceptibility to disease and death [37]. However, frailty models are more appropriate in the context of duration data because of truncation and censoring problems and because of the analytical form of the

likelihood function, which facilitates further numerical calculations. Frailty models can shed much more light on the complicated processes underlying the genetics of longevity.

When parameters of frailty distribution are estimated, the properties of life spans can be studied by using the bivariate survival function, but this is beyond the scope of this paper.

Cause-of-death statistics are taken at face value although there are limitations in time trend comparability with respect to the validity and reliability of cause of death certification and coding and with respect to revisions in the ICD. By describing mortality for a relatively broad diagnostic group like respiratory diseases, the effects of this limitation are lessened.

Misclassification may be a source of bias potentially introduced with the determination of zygosity. Determination of zygosity by means of a questionnaire is a reliable method, with a misclassification rate below 5 per cent when compared with a blood analysis [15, 16]. Since the zygosity has been determined previously and independently of cause of death, no marked bias is introduced with regard to respiratory diseases.

There is one inherent difficulty in this kind of study related to the fact that the data contains only cause of death diagnosed after 1942. About 45 per cent of all twin pairs included in the Danish Twin Registry (especially twins from the older birthcohorts who died at early ages) are lost to the analysis because of unknown cause of death (or unknown zygosity). If genetic factors have a stronger effect at younger ages (as found in Sørensen *et al*. [30]), this loss may lead to an underestimation of genetic effects. Despite this limitation, the present study is one of the largest twin studies in this area with a reasonable number of deaths, resulting from the long follow-up period.

A second limitation of this study is the lack of information about risk factors such as smoking, alcohol intake, diet etc. In interpreting our results, we cannot rule out the possibility that MZ twins shared more environmental risk factors than the DZ twins. For example, MZ twins are more concordant with respect to their smoking behaviour than DZ twins [38]. This greater similarity of experience could be important if these environmental features are predictive of the trait under study and could thus inflate estimates of the genetic influence.

The aetiology of respiratory diseases cannot differ between twins and singletons if the results from twin studies are to be applied to the general population. The fetal origins hypothesis has led to concerns about the validity of the classic twin model. It states that the risk of adult mortality is heightened by retardation in intrauterine growth. The association between low birth weight and asthma [39] and reduced adult forced expiratory volume in one second (FEV_1) [40] support this hypothesis with respect to respiratory diseases. Indeed, twins are, on average, lighter than singletons at birth and may, therefore, have an increased risk of acquiring respiratory diseases in later life. However, comparisons of mortality rates with respect to respiratory diseases show no differences between twins and the general population. Consequently, it seems reasonable to extend the results of the present study to the general population.

Our study suggests that there is a substantial genetic component in frailty to all-cause mortality for both sexes and a modest genetic component in frailty to respiratory disease mortality in females.

ACKNOWLEDGEMENTS

This research was partly supported by NIH/NIA grant 7PO1 AG08761-09. The authors wish to thank Silvia Leek and Karl Brehmer for help in preparing this paper for publication. The presentation was greatly improved by valuable and detailed comments of a referee.

REFERENCES

1. Schächter F, Faure-Delanef, L, Guenot F, Rouger H, Froguel P, Lesueur-Ginot L, Cohen D. Genetic associations with human longevity at the APOE ACE loci. *Nature Genetics* 1994; **6**:29–32.
2. Galinsky D, Tysoe C, Brayne CE, Easton DF, Huppert FA, Dening TR, Paykel ES, Rubinsztein DC. Analysis of the apo E/apo C-I, angiotensin converting enzyme and methylenetetrahydrofolate reductase genes as candidates affecting human longevity. *Atherosclerosis* 1997; **129**:177–183.
3. Yashin AI, Vaupel JW, Andreev KF, Tan Q, Iachine IA, Carotenuto L, de Benedictis G, Bonafe M, Valensin S, Franceschi C. Combining genetic and demographic information in population studies of ageing and longevity. *Journal of Epidemiology and Biostatistics* 1998; **3**:289–294.
4. Cohen BH. Family pattern of mortality and life-span. *The Quarterly Review of Biology* 1964; **39**:130–181.
5. Wyshak G. Fertility and longevity in twins, sibs, and parents of twins. *Social Biology* 1978; **25**:315–330.
6. Herskind AM, McGue M, Holm NV, Sørensen TIA, Harvald B, Vaupel JW. The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900. *Human Genetics* 1996; **97**:319–323.
7. Ljungquist B, Berg S, Lanke J, McClearn G, Pedersen N. The effect of genetic factors for longevity: a comparison of identical and fraternal twins in the Swedish Twin Registry. *Journal of Gerontology* 1998; **53A**:M441–M446.
8. McGue M, Vaupel JW, Holm NV, Harvald B. Longevity is moderately heritable in a sample of Danish twin pairs born 1870–1880. *Journal of Gerontology* 1993; **48**:B237–B244.
9. Hrubec Z, Neel JV. Familial factors in early deaths: twins followed 30 years to ages 51–61 in 1978. *Human Genetics* 1981; **59**:39–46.
10. Jarvik L, Falek A, Kallmann FJ, Lorge I. Survival trends in a senescent twin population. *American Journal of Human Genetics* 1960; **12**:170–179.
11. Yashin AI, Vaupel JW, Iachine I. Correlated individual frailty: an advantageous approach to survival analysis of bivariate data. *Mathematical Population Studies* 1995; **5**:145–159.
12. Yashin AI, Iachine I. Genetic analysis of durations: correlated frailty model applied to the survival of Danish twins. *Genetic Epidemiology* 1995; **12**:529–538.
13. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* 1997; **349**:1498–1504.
14. Armstrong GL, Conn LA, Pinner RW. Trends in infectious disease mortality in the United States during the 20th century. *Journal of the American Medical Association* 1999; **281**:61–66.
15. Lykken DT. The diagnosis of zygosity in twins. *Behavior Genetics* 1978; **8**:437–473.
16. Holm NV. The use of twin studies to investigate causes of diseases with complex etiology with a focus on cancer (in Danish). *Ph.D. Thesis*, Odense University, Odense, 1983.
17. Hauge M. The Danish Twin Register. In *Prospective Longitudinal Research*, Mednich SA, Baert AE, Bachmann BP (eds). Oxford Medical: Oxford, 1968; 217–222.
18. Juel K, Helweg-Larsen K. The Danish registers of causes of death. *Danish Medical Bulletin* 1999; **46**:354–357.
19. Christensen K, Vaupel JW, Holm NV, Yashin AI. Mortality among twins after age 6: fetal origins hypothesis versus twin method. *British Medical Journal* 1995; **310**:432–436.
20. Taeger D, Sun Y, Keil U, Straif K. A stand-alone windows application for computing exact person years, standardized mortality ratios and confidence intervals in epidemiological studies. *Epidemiology* 2000; **11**: 607–608.
21. McGue M. When assessing the twin concordance, use the probandwise not the pairwise rate. *Schizophrenia Bulletin* 1992; **18**:171–176.
22. Nielsen GG, Gill RD, Andersen PK, Sørensen TIA. A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics* 1992; **19**:25–43.
23. Hougaard P. *Analysis of Multivariate Survival Data*. Springer: New York, 2000.
24. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 1979; **16**:439–454.
25. Neale MC, Cardon LR. *Methodology for Genetic Studies of Twins and Families*. Kluwer, Dordrecht, 1992.
26. Yashin AI, Iachine IA. Environment determines 50% of variability in individual frailty: results from Danish twin study. *Research Report. Population Studies of Aging*, vol. 10, Odense University, Denmark, 1994.
27. Statistical Package for the Social Sciences (SPSS). SPSS, Inc.: Chicago, 1989–1997.
28. GAUSS. Aptech Systems, Inc.: Maple Valley WA, 1984–1996.
29. Yashin AI, Iachine I. How long can humans live? Lower bound for biological limit of human longevity calculated from Danish twin data using correlated frailty model. *Mechanisms of Aging and Development* 1995; **80**: 147–169.
30. Sørensen TIA, Nielsen GG, Anderson PK, Teasdale TW. Genetic and environmental influences on premature death in adult adoptees. *New England Journal of Medicine* 1988; **318**:727–732.
31. Yashin AI, Iachine I. What difference does the dependence between durations make? Insights for population studies of aging. *Lifetime Data Analysis* 1999; **5**:5–22.
32. Skadhauge LR, Christensen K, Kyvik KO, Sigsgaard T. Genetic and environmental influence on asthma: a population based study of 11,688 Danish twin pairs. *European Respiratory Journal* 1999; **13**:8–14.

33. Duffy DL, Martin NG, Battistutta D, Hopper JL, Mathews JD. Genetics of asthma and hay fever in Australian twins. *American Review of Respiratory Diseases* 1990; **142**:1351–1358.
34. Harris JR, Magnus P, Samuelsen SO, Tambs K. No evidence for effects of family environment on asthma. A retrospective study of Norwegian twins. *American Journal of Respiratory and Critical Care Medicine* 1997; **156**:43–49.
35. Koppelman GH, Los H, Postma DS. Genetic and environment in asthma: the answer of twin studies. *European Respiratory Journal* 1999; **13**:2–4.
36. Nieminen MM, Kaprio J, Koskenvuo M. A population-based study of bronchial asthma in adult twin pairs. *Chest* 1991; **100**:70–75.
37. Falconer DS, Mackay TF. *Introduction to Quantitative Genetics*. Longman: London, 1996.
38. Carmelli D, Swan GE, Robinette CD, Fabsitz R. Genetic influence on smoking—a study of male twins. *New England Journal of Medicine* 1992; **327**:829–833.
39. Schwartz J, Gold D, Dockery DW, Weiss ST, Speizer FE. Predictors of asthma persistent wheeze in a national sample of children in the United States. Association with social class, perinatal events, and race. *American Review of Respiratory Diseases* 1990; **142**:555–562.
40. Barker DJ, Godfrey KM, Fall C, Osmond C, Winter PD, Shaheen SO. Relation of birth weight and childhood respiratory infection to adult lung function and death from chronic obstructive airways diseases. *British Medical Journal* 1991; **303**:671–675.

