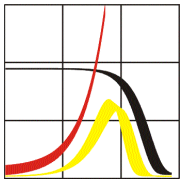

Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research
Doberaner Strasse 114 · D-18057 Rostock · GERMANY
Tel +49 (0) 3 81 20 81 - 0; Fax +49 (0) 3 81 20 81 - 202;
http://www.demogr.mpg.de

MPIDR WORKING PAPER WP 2002-003
JANUARY 2002 (REVISED SEPTEMBER 2002)

The influence of smoking and BMI on heritability in susceptibility to coronary heart disease

Andreas Wienke (wienke@demogr.mpg.de)
Anne M. Herskind
Kaare Christensen (kchristensen@health.sdu.dk)
Axel Skytthe (askytthe@health.sdu.dk)
Anatoli I. Yashin (yashin@demogr.mpg.de)

Abstract

Cause-specific mortality data on Danish monozygotic (MZ) and dizygotic (DZ) twins are used to analyze the influence of smoking and body mass index (BMI) on heritability estimates of susceptibility to coronary heart disease (CHD). The sample includes 1209 like-sexed twin pairs born between 1890 and 1920, where both individuals were still alive and answered a questionnaire, including information about smoking, height and weight, in 1966. The analysis was conducted with both sexes pooled due to the relatively small number of twin pairs. Follow-up was conducted from 1 January 1966 to 31 December 1993. We use the correlated gamma-frailty model with observed covariates for the genetic analysis of frailty to account for censoring and truncation in the lifetime data. During the follow-up, 1437 deaths occurred, including 435 deaths due to CHD. Proportions of variance of frailty attributable to genetic and environmental factors were analyzed using the structural equation model approach. Different 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 model without covariates, heritability of frailty to CHD was found to be 0.45 (0.11). This result changes only slightly to 0.54 (0.16) after controlling for smoking and BMI. This analysis underlines the existence of a substantial genetic influence on individual frailty associated with mortality caused by CHD. No evidence for common genetic factors acting on smoking, BMI, and susceptibility to CHD are found which indicates that the association between smoking and susceptibility to CHD and BMI and susceptibility to CHD is not confounded by common genetic factors.

1 Introduction

Twin studies is one of the most widely used methods for quantifying the influence of genetic and environmental factors on specific diseases. In the case of binary traits (where disease is either present or not), concordance analysis provides a powerful and widely accepted method in genetic epidemiology. Concordance rates are easily to calculate and allow a clear interpretation (McGue[1], Gatz et al.[2]). In practical applications, time-to-event data (time of onset of disease, age at death) is often available, but usually in a truncated and/or censored form. Censoring of bivariate observations can be a complex problem, as either or both individuals of a pair may be subject to censoring, and the censoring times need not to be the same for both individuals. Additional problems arise through bivariate truncation, which implies a non-random selection of the study population from the total twin population. Furthermore, in many cases covariates are available. Unfortunately, it is difficult to manage truncated and censored time-to-event data and covariates within the context of concordance analysis. A large part of the motivation for the methodology of this article is exploring the potential for censored and truncated data and the inclusion of measured covariates. One key question here is whether the inclusion of covariates changes the heritability estimates of susceptibility to

the disease under study.

We aim to study bivariate survival times X_1, X_2 which depend, via a proportional hazards models, on unobserved variables Z_1, Z_2 (called frailties). That is, we seek to explain an association between two (non-negative) survival times. They have a continuous joint distribution through their common dependence upon unobserved random variables. Such models are particularly convenient in the context of survival data and they stem from the (univariate) concept of frailty introduced by Vaupel et al.[3]. The univariate frailty model extends immediately to frailty models with a bivariate survival function $S(x_1, x_2) = P(X_1 > x_1, X_2 > x_2)$. The bivariate frailty model asserts that X_1 and X_2 are conditionally independent, given Z_1, Z_2, and that:

$$S(x_1, x_2) = \mathbf{E}e^{-Z_1 H_{01}(x_1) - Z_2 H_{02}(x_2)}, \tag{1}$$

for some cumulative baseline hazard functions H_{01} and H_{02}. We will give more detailed information about this model later on.

We want to apply the bivariate frailty model to twin data on mortality due to coronary heart disease (CHD). The role of family aggregation of CHD is well established. Questions about the nature of the genetic effects (additive versus non-additive) are addressable. Former studies from the Danish[4] and Swedish[5,6] twin registries found a genetic component in the risk of death from coronary heart disease. Recent studies[7,8] used approaches from survival analysis to account for truncation and censoring present in the data. Such kinds of methods have to combined with methods from genetic epidemiology. Heterogeneity of individuals with susceptibility to CHD as well as important covariates is included in the present model. For such a combined analysis, we apply the correlated gamma-frailty model with observed covariates[9,10] in the present paper, which takes into account the dependence of life spans of relatives (twins). This allows the estimation of the effect of genetic factors in susceptibility to CHD and the evaluation as to what extend smoking, BMI, and susceptibility to CHD are all influenced by common genetic factors. This approach enables the combination of data on the age at death with data on cause of death, smoking, and BMI; and further, to deal with truncated and censored observations. For each individual we assume two independent underlying competing risks of latent times (lifetime with respect to death due CHD and lifetime with respect to death due to all other diseases (including censoring)). In addition, we assume that these competing risks are independent. An extension of the model to more than two competing risks is possible, but beyond the scope of the present paper. We empirically demonstrate the advantages of the model in the statistical analysis of lifetime data from Danish twins, which were already used in Herskind et al.[11], but now with a special focus on mortality caused by CHD and applying a frailty model. The model allows checking hypothesis about genetic confounding in the relationship between the covariates (smoking and BMI) and susceptibility to CHD. The concept of genetic confounding was originally introduced by R.A. Fisher, who suggested that the association between smoking and lung cancer

was caused by common genetic factors[12]. For references related to the discussion in this field, see Herskind et al.[11].

2 Material and methods

Mortality data of twins were provided by the Danish Twin Registry, founded in 1954 as the world's first nation-wide twin registry. This population-based registry includes twins born in Denmark during the period 1870-1910 and all like-sex pairs born between 1911 and 1930. For detailed information about the Danish Twin Registry, see Hauge[13].

In 1966, a questionnaire including questions about smoking, height, and weight was mailed to all twins born 1890-1920 who were alive and traceable on 1 January 1966. 3709 individuals answered the questionnaire (response rate 65 %). Excluded from the study were 813 twins with non-responding partners, four pairs with unknown zygosity and 212 pairs with incomplete or uncertain information on height and weight. 23 pairs were excluded because of incomplete information about cause of death, resulting in a study population of 1209 twin pairs.

Individuals were followed from 1 January 1966 to 31 December 1993. Those persons identified as deceased after that date are classified for our purposes as 'living'. At the end of follow-up, approximately 40 % of the twins were still alive, resulting in right censored data. Altogether, there are 210 male monozygotic twin pairs and 316 dizygotic twin pairs, 273 female monozygotic twin pairs and 410 dizygotic twin pairs. In addition to age at death, there is also information on cause of death available for all individuals who died during the follow-up. For the present study, only the underlying cause of death was considered. Detailed information about death status, gender, zygosity, smoking, and BMI of the study population is given in Table 1 and 2.

status	males		females	
	MZ twins	DZ twins	MZ twins	DZ twins
deaths				
–CHD	96	153	76	110
–other	206	280	204	312
all causes	302	433	280	422
alive	118	199	266	398
total	420	632	546	820

Table 1: Study population (number of individuals) by gender, zygosity and cause of death

	Non-smokers	Other (ex-smokers etc.)	pipe/cigar smokers	cigarette smokers	Total
males					
BMI < 22	11	12	45	31	99 (9.4%)
BMI 22- 28	78	141	333	208	760 (72.2%)
BMI > 28	24	35	91	43	193 (18.3%)
Total	113 (10.7%)	188 (17.9%)	469 (44.6%)	282 (26.8%)	
females					
BMI < 22	105	46	47	99	297 (21.7%)
BMI 22- 28	350	138	134	171	793 (58.1%)
BMI > 28	157	40	37	42	276 (20.2%)
Total	612 (44.8%)	224 (16.4%)	218 (16.0%)	312 (22.8%)	

Table 2: Study population (number of individuals) by sex, BMI and smoking

2.1 Mortality

After the age of six, 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 monozygotic twins are close to those of dizygotic twins for both sexes[14]. Recent papers dealing with twin cohorts born during the period of 1870 - 1930 found similar mortality patterns for Danish twins and the general Danish population with respect to CHD[8,15]. This similarity suggest that it is possible to generalize genetic results from survival analysis of twins to the total population with respect to mortality due to CHD.
For the present report, CHD is grouped as ICD 420 in the sixth and seventh revision and as ICD 410 - 414 in the eighth ICD revision.

2.2 Statistical methods

Univariate lifetime models cannot capture the association between the life spans of related individuals like twins. Consequently, bivariate distributions of dependent lifetimes are necessary. For genetic analysis of time-to-event data, associations between durations are needed. In this paper, we want to analyze genetic and environmental factors acting on susceptibility (frailty) to mortality due to CHD when controlled for smoking and BMI. The correlated gamma-frailty model with observable covariates can be used to fit bivariate lifetime data and provide a specific parameter for correlation of frailty to death. The interesting point here is, that individual frailties in twin pairs could not be observed, but their correlation can be estimated by application of the correlated gamma-frailty model. Now we want to make more specific assumptions about the structure of the lifetimes. To include heterogeneity in our model, we assume a correlated gamma-frailty model[16,17]. Let Z_i $(i = 1, 2)$ be the frailties, and u_i $(i = 1, 2)$ vectors of observable covariates of the two individuals of a twin pair. Assume that their individual hazards are represented by the

proportional hazards model $\mu(x, Z_i, u_i) = Z_i \mu_0(x) e^{\beta u_i}$ $(i = 1, 2)$ with a baseline hazard function $\mu_0(x)$ describing the risk of dying as a function of age and β denotes the vector of regression parameters. Let the lifetimes of the two twin partners be conditionally independent given their frailties Z_1 and Z_2. Additionally, 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 $Y_0 \sim \Gamma(k_0, \lambda)$, $Y_1 \sim \Gamma(k_1, \lambda)$ and $Y_2 \sim \Gamma(k_2, \lambda)$. Here k_0, k_1, k_2, λ are non-negative parameters and $\Gamma(k, \lambda)$ denotes a Gamma distribution with parameters k and λ. 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 with respect to twins, because there is no reason to assume different distributions of frailty in twin partners. We employ the standard assumption that the mean frailty of individuals is one (at the beginning of the follow-up), which means that $\mathbf{E}Z_i = \frac{k_0 + k_i}{\lambda} = 1$ $(i = 1, 2)$. 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.[17]. The unconditional bivariate survival function is given by:

$$S(x_1, x_2|u_1, u_2) = S(x_1|u_1)^{1-\rho} S(x_2|u_2)^{1-\rho} (S(x_1|u_1)^{-\sigma^2} + S(x_2|u_2)^{-\sigma^2} - 1)^{-\frac{\rho}{\sigma^2}}, \qquad (2)$$

where $S(x|u)$ denotes the marginal univariate survival function, assumed to be equal for both partners in a twin pair. Using a parametric approach we fitted a Gamma-Gompertz model to the data, e.g. $S(x|u) = \left(1 + [(1 + s^2 \frac{\alpha}{\gamma}(e^{\gamma x} - 1))^{\frac{\sigma^2}{s^2}} - 1] e^{\beta u}\right)^{-\frac{1}{s^2}}$, where $\alpha, \beta, s^2, \sigma^2, \rho, \gamma$ are parameters to be estimated.

Let $(X_{11}, X_{12}, U_{11}, U_{12}), \ldots, (X_{n1}, X_{n2}, U_{n1}, U_{n2})$ be independent and identically distributed (i.i.d.) lifetimes and observable covariates. The lifetimes (X_{i1}, X_{i2}) are assumed to be independently censored from the right by i.i.d. pairs of non-negative random variables $(C_{11}, C_{12}), \ldots, (C_{n1}, C_{n2})$, which are independent of the (X_{i1}, X_{i2}). Thus, instead of $(X_{i1}, X_{i2}, U_{i1}, U_{i2})$ we only observe

$$(T_{i1}, T_{i2}, \Delta_{i1}, \Delta_{i2}, U_{i1}, U_{i2}) \qquad (3)$$

with $T_{ij} = \min\{X_{ij}, C_{ij}\}$, $\Delta_{ij} = 1(X_{ij} \leq C_{ij})$ $(i = 1, \ldots, n;\ j = 1, 2)$, where $1(\cdot)$ denotes the indicator function of the event in the brackets. Let us assume that the lifetimes follow a distribution (dependent on observable covariates U_1, U_2) given by the bivariate survival function $S(x_1, x_2|u_1, u_2) = P(X_{i1} > x_1, X_{i2} > x_2|u_1, u_2)$, and denotes by $C(c_1, c_2) = P(C_{i1} > c_1, C_{i2} > c_2)$ the survival function of censoring times. Hence, the survival function of the four-dimensional latent times is of the form:

$$S(x_1, c_1, x_2, c_2|u_1, u_2) = S(x_1, x_2|u_1, u_2)C(c_1, c_2). \tag{4}$$

Starting from this model, we are able to derive the likelihood function of the data given by (3):

$$\begin{aligned} L(t_1, t_2, \delta_1, \delta_2, u_1, u_2) &= \delta_1\delta_2 S_{t_1t_2}(t_1, t_2|u_1, u_2) - \delta_1(1-\delta_2)S_{t_1}(t_1, t_2, u_1, u_2) \\ &\quad - (1-\delta_1)\delta_2 S_{t_2}(t_1, t_2|u_1, u_2) + (1-\delta_1)(1-\delta_2)S(t_1, t_2|u_1, u_2) \end{aligned} \tag{5}$$

with $(t_1, t_2, \delta_1, \delta_2, u_1, u_2)$ as a realisation of the random vector $(T_1, T_2, \Delta_1, \Delta_2, U_1, U_2)$. Partial derivatives of the marginal survival functions are given by $S_{t_i}(t_1, t_2) = \frac{\partial S(t_1, t_2)}{\partial t_i}$ $(i = 1, 2)$ and $S_{t_1t_2}(t_1, t_2) = \frac{\partial S(t_1, t_2)}{\partial t_1 \partial t_2}$. Because of the independence assumption between lifetimes (X_{i1}, X_{i2}) and censoring times (C_{i1}, C_{i2}) the distribution of the censoring times does not enter the likelihood function.

As mentioned above, the twin pair data set used is not randomly selected from the total twin population. Since both members of a twin pair had to be still alive on 1 January 1966, the survival times in the data set are sampled from specific conditional distributions. If a twin pair was born in year y (where y=1880, ... ,1920), the condition of survival of both twins until the year 1966 implies that both twins had to survive until the age of 1966-y in order to be included in the sample. If the survival times are denoted by X_1 and X_2 with survival function $S(x_1, x_2|u_1, u_2)$, then the conditional survival function for a twin pair born in year y is:

$$S(x_1, x_2|u_1, u_2, X_1 > 1966 - y, X_2 > 1966 - y) = \frac{S(x_1, x_2|u_1, u_2)}{S(1966 - y, 1966 - y|u_1, u_2)}$$

Consequently, the likelihood function of independently left truncated and right censored lifetime data is given by:

$$\begin{aligned} L(t_1, t_2, \delta_1, \delta_2, u_1, u_2, y) &= \frac{\delta_1\delta_2 S_{t_1t_2}(t_1, t_2|u_1, u_2) - \delta_1(1-\delta_2)S_{t_1}(t_1, t_2|u_1, u_2)}{S(1966 - y, 1966 - y|u_1, u_2)} \\ &\quad - \frac{(1-\delta_1)\delta_2 S_{t_2}(t_1, t_2|u_1, u_2) + (1-\delta_1)(1-\delta_2)S(t_1, t_2|u_1, u_2)}{S(1966 - y, 1966 - y|u_1, u_2)} \end{aligned} \tag{6}$$

For a combined analysis of monozygotic and dizygotic twins we include two correlation coefficients, ρ_{MZ} and ρ_{DZ}, respectively. These correlations between monozygotic and dizygotic twins provide information about genetic and environmental influences on frailty within individuals.

2.3 Quantitative genetics of frailty

In twin studies, the intrapair-correlations of the trait under study (here frailty on mortality due to CHD) in monozygotic and dizygotic twin pairs play the key role for analysis of genetic and environmental factors. Using these coefficients, five standard genetic models of frailty are fitted to the data that corresponds to five different assumptions about its structure. Resemblance in twins is (completely for monozygotic twins and partly for dizygotic twins) caused by two factors: additive genetic factors (A) and shared environmental factors (C). Non-shared environment is (completely for monozygotic twins and partly for dizygotic twins) responsible for intra-pair differences in twins. From the estimation point of view, three parameters could be included into the model simultaneously, because there are data about two different groups of relatives (monozygotic and dizygotic twins). Models that are more complex need data about additional groups of relatives. Each additional group of relatives allow for an additional parameter in the model. The following biometric models were fitted to the data: ACE, ADE, AE, DE and CE. In these notations, an ACE model refers to the decomposition of frailty Z=A+C+E. ADE, AE, DE and CE models are defined similarly. We use the small letters a^2, c^2, e^2, d^2 to refer to the respective proportions of variance. For example, the relation $1 = a^2 + c^2 + e^2$ corresponds to the decomposition of variance in the ACE model of frailty. Standard assumptions about of the quantitative genetics yields in the following relations:

$$\begin{aligned} \rho_{MZ} &= a^2 + d^2 + c^2 \\ \rho_{DZ} &= 0.5a^2 + 0.25d^2 + c^2 \\ 1 &= a^2 + d^2 + c^2 + e^2. \end{aligned} \tag{7}$$

For detailed information about these genetic models, and deriving the upper equations, see Neale and Cardon[18]. Note that the AE model and the CE model are not nested. Consequently, the likelihood ratio test could not be used to define the model with the best fit to the data. The Akaike Information Criterion (AIC)[19] is used to compare non-nested models.

To combine the approach of quantitative genetics with the methods of survival analysis we used the correlated gamma-frailty model with genetic and environmental components of frailty. In this approach, the decompositions in (7) must be substituted into survival model (2). This model must be used to estimate the parameters a^2, d^2, c^2, e^2 by the maximum likelihood method directly. Analysis was made using standard statistical software packages SPSS[20] and GAUSS[21].

3 Results

Because not all models are nested, the likelihood ratio test can not be applied to compare all five biometric models. Applying the correlated gamma-frailty model with and without observed covariates the Akaike Information Criterion prefers the AE model. Using this model, heritability changes from 0.45 (0.11) without covariates, to 0.54 (0.16) with covariates (see Table 3). Standard errors for the ACE model are not shown in Table 3 since $c^2 = 0$ is the boundary of the parameter space.

model	σ	a^2	d^2	c^2	e^2	β_1	β_2	β_3	β_4	β_5	AIC
ACE	2.78	0.45		0.00	0.55						
	(-)	(-)		(-)	(-)						4566.1
ACE	1.99	0.54		0.00	0.46	0.14	0.29	1.00	0.92	0.81	
	(-)	(-)		(-)	(-)	(-)	(-)	(-)	(-)	(-)	4549.5
AE	2.78	0.45			0.55						
	(0.72)	(0.11)			(0.11)						4564.1
AE	1.99	0.54			0.46	0.14	0.29	1.00	0.92	0.81	
	(0.56)	(0.16)			(0.16)	(0.22)	(0.22)	(0.27)	(0.25)	(0.27)	4547.5
ADE	2.78	0.44	0.01		0.55						
	(0.80)	(1.19)	(1.40)		(0.25)						4566.1
ADE	1.95	0.31	0.28		0.42	0.14	0.29	0.99	0.91	0.80	
	(0.56)	(0.44)	(0.51)		(0.19)	(0.22)	(0.21)	(0.27)	(0.25)	(0.27)	4549.2
DE	2.68		0.50		0.50						
	(0.76)		(0.13)		(0.13)						4565.9
DE	1.88		0.62		0.38	0.14	0.29	0.99	0.90	0.79	
	(0.54)		(0.19)		(0.19)	(0.21)	(0.21)	(0.26)	(0.24)	(0.26)	4547.7
CE	2.96			0.30	0.70						
	(0.76)			(0.08)	(0.08)						4567.4
CE	2.11			0.35	0.65	0.14	0.27	1.01	0.94	0.84	
	(0.62)			(0.11)	(0.11)	(0.23)	(0.22)	(0.29)	(0.26)	(0.29)	4552.0

Table 3: Estimates of the components of variance in frailty to mortality from CHD(n=1209). σ^2 - variance of frailty, a^2 - additive genetic effects, d^2 - genetic effects as a result of dominance, c^2 - common environment, e^2 - non-shared (individual) environment (including measurement errors), AIC - Akaike Information Criterion, β_1 - BMI $< 22\,Kg/m^2$, β_2 - BMI $> 28\,Kg/m^2$, β_3 - cigarette smokers, β_4 - pipe/cigar smokers, β_5 - former smokers

Using the best fitting AE model, the likelihood ratio test does not indicate any significant influence of BMI on CHD mortality ($\beta_1 = 0.14\,(0.22)$ and $\beta_2 = 0.29\,(0.22)$). Here β_1 and β_2 describe individuals with BMI less than $22kg/m^2$ and more than $28kg/m^2$, respectively. The reference group are individuals with BMI between $22kg/m^2$ and $28kg/m^2$. Smoking shows a significant influence on CHD mortality ($\beta_3 = 1.00\,(0.27)$, $\beta_4 = 0.92\,(0.25)$ and $\beta_5 = 0.81\,(0.27)$). Here β_3, β_4 and β_5 denote cigarette smokers, pipe and cigar smokers and former smokers, respectively. The reference group is the nonsmokers.

4 Discussion

The presented method in this article with its suitability for censored and truncated data and the possibility to include observed covariates allows to overcome the well-known drawbacks of the traditional concordance analysis in twin studies with time-to-event data. An important question arising in genetic analysis of models with observed covariates (in our case smoking and BMI) is whether genes that are responsible for variation in observed covariates also contribute to a variation in susceptibility to CHD. In that case traditional biometrical methods of regression analysis can led to spurious effects of covariates. In extrem cases not covariates, but common genes may be responsible for variation in life span.

Both smoking and BMI are influenced by genes with heritability estimates $0.35 - 0.75$ (smoking) and $0.5 - 0.8$ (BMI) (Bouchard[22], Heath and Madden[23], Herskind et al.[24]). However, whether common genes influence these phenotypic traits, as well as susceptibility to CHD, is an open question. In 1958 R.A. Fisher[12] suggested that the association between smoking and lung cancer is spurious and reflects only the circumstance that the same genes influence both smoking habits and lung cancer. This was the starting point for a long debate on genetic confounding.

The main result of the present paper was that the inclusion of smoking and BMI do not cause any substantial decrease in the heritability estimates. Hence, no evidence was found for common genetic factors acting on smoking and susceptibility to CHD or BMI and susceptibility to CHD. This study confirms the earlier finding that the genetic influence on susceptibility to CHD is not mediated trough genetic influence on smoking and BMI.

As expected the inclusion of observable covariates decreases the heterogeneity in the population, which can be seen in the decline of the variance of the frailty from $\sigma = 2.78$ in the model without covariates to $\sigma = 1.99$ in the model with covariates. This makes clear that frailty is not a phenotypic trait. Frailty depends on the model, it describes factors not included in the model.

When observed covariates are included in the model, the relative importance of environmental factors (shared environment, C, and non-shared environment, E) is reduced, leading to an increase in the heritability estimates observed in the present analysis. However, this increase is negligible with respect to the calculated standard errors of these estimates. Genetic confounding would lead to a decrease in heritability estimates because in that case genetic factors contribute predominantly to the observed covariates rather than to unobserved covariates included in the frailty.

The analysis underlines the importance of genetic factors on individual susceptibility to CHD. Otherwise, the heritability estimate in our study (0.45 for both sexes combined) is lower than those found in a previous analysis of an extension of the presented data set (without covariates) with heritability estimates of 0.53 and 0.58 for males and females, respectively (Wienke et al.[8]). The lower estimates in the subsample analyzed in

the present paper may be a consequence of a decline in the heritability of CHD with increasing age as found in Marenberg et al.[6] and Zdravkovic et al.[7]. The twin population in our study is much older because time of truncation was 1966 compared with 1943 in the paper by Wienke et al.[8]. Furthermore, the youngest cohorts (birth years 1921-1930) are not included in the present analysis.
The proposed method allows to handle censored and truncated time-to-event data of related individuals in the case of observed covariates. The hypothesis of genetic confounding can be checked and the influence of observable covariates on heritability can be analyzed.

Acknowledgments

The authors wish to thank the Danish Twin Register for providing the twin data and Susann Backer for help in preparing the paper for publication. The research was partly supported by NIH/NIA grant 7PO1AG08761-09.

References

[1] McGue M. When Assessing Twin Concordance, Use the Probandwise Not the Pairwise Rate. *Schizophrenia Bulletin* 1992; 18: 171 - 176.

[2] Gatz M, Pedersen NL, Crowe M, Fiske A. Defining discordance in twin studies of risk and protective factors for late life disorders. *Twin Research* 2000; 3: 159 - 64.

[3] Vaupel JW, Manton KG, Stallard E. The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. *Demography* 1979; 16: 439 - 454.

[4] Harvald B, Hauge M. Coronary occlusion in twins. *Acta Geneticae Medicae et Gemellologiae* 1970; 19: 248 - 250.

[5] de Faire U, Friberg L, Lundman T. Concordance for mortality with special reference to ischaemic heart disease and cerebrovascular disease: a study on the Swedish Twin Registry. *Preventive Medicine* 1975; 4: 509 - 517.

[6] Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. *The New England Journal of Medicine* 1994; 330: 1041 - 1046.

[7] Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, de Faire U. Heritability of death from coronary heart disease: 36 years follow-up of 20,966 Swedish twins. *Journal of Internal Medicine, forthcoming*

[8] Wienke A, Holm NV, Skytthe A, Yashin AI. The heritability of mortality due to heart diseases: a correlated frailty model applied to Danish twins. *Twin Research* 2001; 4: 266 - 274.

[9] Yashin AI, Manton KG, Iachine IA. Genetic and environmental factors in duration studies: multivariate frailty models and estimation strategies. *Journal of Epidemiology and Biostatistics* 1996; 1: 115 - 120.

[10] Yashin AI, Iachine IA. How frailty models can be used for evaluating longevity limits: Taking advantage of an interdisciplinary approach. *Demography* 1997; 34: 31 - 48.

[11] Herskind AM, McGue M, Iachine IA, Holm N, Sørensen TIA, Harvald B, Vaupel JW. Untangling genetic influences on smoking, body mass index and longevity: a multivariate study of 2464 Danish twins followed for 28 years. *Human Genetics* 1996; 98: 467 - 475.

[12] Fisher RA. Cancer and smoking. *Nature* 1958; 182: 596.

[13] Hauge M. The Danish Twin Register. In: Mednich SA, Baert AE, Bachmann BP. eds. *Prospective longitudinal research.* Oxford: Oxford Medical Press, 1968: 217 - 222.

[14] 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.

[15] Christensen K, Wienke A, Skytthe A, Holm NV, Vaupel JW, Yashin AI. Cardiovascular mortality in twins and the fetal origins hypothesis. Twin Research 2001; 5: 344 - 349

[16] Yashin AI, Iachine IA. How long can humans live? Lower bound for biological limit of human longevity calculated from Danish twin data using correlated frailty model. *Mechanisms of Ageing and Development* 1995; 80: 147 - 169.

[17] Yashin AI, Vaupel JW, Iachine IA. Correlated individual frailty: An advantageous approach to survival analysis of bivariate data. *Mathematical Population Studies* 1995; 5: 145 - 159.

[18] Neale MC, Cardon LR. *Methodology for genetic studies of twins and families.* Dordrecht, Kluwer, 1992.

[19] Akaike H. Factor analysis and AIC. *Psychometrica* 1987; 52: 317 - 322.

[20] GAUSS. Aptech Systems, Inc. Maple Valley WA 1984-1996.

[21] Statistical Package for the Social Sciences (SPSS). Chicago: SPSS, Inc. 1989-1997.

[22] Bouchard C., (1994). Genetics of obesity: Overview and research directions. In: Claude Bouchard (ed) The genetics of obesity. 1st edn. CRC Press, Boca Raton, pp. 223-233.

[23] Heath AC., Madden PAF (1995). Genetic influences on smoking behavior. In: Turner JR., Cardon LR., Hewitt JK., (eds). Behavior: Genetic approaches in behavioral medicine. Plenum Press, New York, London, pp. 45-63.

[24] Herskind AM., McGue. M, Sorensen TIA., Harvald B. (1996). Sex and age specific assessment of genetic and environmental influences on body mass index in twins. International Journal of Obesity 20: 106-113.