

Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research
Konrad-Zuse-Strasse 1 · 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 2003-032
SEPTEMBER 2003

Frailty Models

Andreas Wienke

This working paper has been approved for release by: Anatoli Yashin (yashin@demogr.mpg.de),
Head of the Laboratory of Advanced Statistical Methods.

© Copyright is held by the authors.

Working papers of the Max Planck Institute for Demographic Research receive only limited review.
Views or opinions expressed in working papers are attributable to the authors and do not necessarily
reflect those of the Institute.

Introduction

The notion of frailty provides a convenient way to introduce random effects, association and unobserved heterogeneity into models for survival data. In its simplest form, a frailty is an unobserved random proportionality factor that modifies the hazard function of an individual, or of related individuals. In essence, the frailty concept goes back to work of Greenwood and Yule (1920) on "accident proneness". The term frailty itself was introduced by Vaupel et al. (1979) in univariate survival models and the model was substantially promoted by its application to multivariate survival data in a seminal paper by Clayton (1978) (without using the notion "frailty") on chronic disease incidence in families.

Frailty models are extensions of the proportional hazards model which is best known as the Cox model (Cox, 1972), the most popular model in survival analysis. Normally, in most clinical applications, survival analysis implicitly assumes a homogenous population to be studied. This means that all individuals sampled into that study are subject in principle under the same risk (e.g., risk of death, risk of disease recurrence). In many applications, the study population can not be assumed to be homogeneous but must be considered as a heterogeneous sample, i.e. a mixture of individuals with different hazards. For example, in many cases it is impossible to measure all relevant covariates related to the disease of interest, sometimes because of economical reasons, sometimes the importance of some covariates is still unknown. The frailty approach is a statistical modelling concept which aims to account for heterogeneity, caused by unmeasured covariates. In statistical terms, a frailty model is a random effect model for time-to-event data, where the random effect (the frailty) has a multiplicative effect on the baseline hazard function. One can distinguish two broad classes of frailty models:

- (1) models with an univariate survival time as endpoint and
- (2) models which describe multivariate survival endpoints (e.g; competing risks, recurrence of events in the same individual, occurrence of a disease in relatives).

In the first case, an univariate (independent) lifetime is used to describe the influence of unobserved covariates in a proportional hazards model (heterogeneity). The variability of survival data is split into a part that depends on risk factors, and is therefore theoretically predictable, and a part that is initially unpredictable, even when all relevant information is known. A separation of these two sources of variability has the advantage that heterogeneity can explain some unexpected results or give an alternative interpretation of some results, for example, crossing-over effects or convergence of hazard functions of two different treatment arms (see Manton and Stallard (1981)) or leveling-off effects - that means the decline in the increase of mortality rates, which could result in a hazard function at old ages parallel to the x-axis (Aalen and Tretli (1999)). More interesting, however, is the second case when multivariate survival times are considered. There one aims to account for the dependence in clustered event times, for example in the lifetimes of patients in study centers in a multi-center clinical trial, caused by center-specific conditions (see Andersen et al. (1999)). A natural way to model dependence of clustered event times is through the introduction of a cluster-specific random effect - the frailty. This random effect explains the dependence in the sense that had we known the frailty, the events would be independent. In other words, the lifetimes are conditional independent, given the frailty. This approach can be used for survival times of related individuals like family members or recurrent observations on the same person. Different extensions of univariate frailty models to multivariate models are possible and will be considered below.

We want to explain the key ideas of univariate frailty models by an illustrative example from Aalen and Tretli (1999). The authors analyzed the incidence of testis cancer by means of a frailty model based on data from the Norwegian Cancer Registry collected during 1953-93. The incidence of testicular cancer is greatest among younger men, and then declines from a certain age. The frailty is considered to be established by birth, and due to a mixture of genetic and environmental effects. The idea of the frailty model is that a subgroup of men is particularly susceptible to testicular cancer. This would explain why testis cancer is primarily a disease of

young men. As time goes by the members of the frail group acquire the disease, and at some age this group is more or less exhausted. Then the incidence, computed on the basis of all men at a certain age, will necessarily decline.

Univariate frailty models

The standard situation of the application of survival methods in clinical research projects assumes that a homogeneous population is investigated when subject under different conditions (e.g. experimental treatment and standard treatment). The appropriate survival model then assumes that the survival data of the different patients are independent from each other and that each patient's individual survival time distribution is the same (independent and identically distributed failure times).

This basic presumption implies a homogeneous population. However, in the field of clinical trials one observes in many most practical situations that patients differ substantially. The effect of a drug, a treatment or the influence of various explanatory variables may differ greatly between subgroups of patients. To account for such unobserved heterogeneity in the study population Vaupel et al. (1979) introduced univariate frailty models into survival analysis. The key idea is, that individuals possess different frailties, and that those patients who are most frail will die earlier than the others. Consequently, systematic selection of robust individuals (that means patients with low frailty) takes place. When mortality rates are estimated, one may be interested in how these rates change over time or age. Quite often it is observed that the hazard function (or mortality rate) rises at the beginning, reaches a maximum, and then declines (unimodal intensity) or levels-off at a constant value. The longer the patient lives after manifestation of the disease, the more improved are his or her chances of survival. It is likely that unimodal intensities are often a result of a selection process acting in a heterogeneous population and do not reflect individual

mortality. The population intensity may start to decline simply because the high-risk individuals have already died out. The hazard rate of a given individual might well continue to increase.

If protective factors or risk factors are known, those could be included in the analysis by using the proportional hazards model, which is of the form

$$\mu(t, X) = \mu_0(t) \exp(\beta^T X),$$

where $\mu_0(t)$ denotes the baseline hazard function, assumed to be unique for all individuals in the study population. X is the vector of observed covariates and β the respective vector of regression parameters to be estimated. The mathematical convenience of this model is based on the separation of the effects of aging in the baseline hazard $\mu_0(t)$ from the effects of covariates in the parametric term $\exp(\beta^T X)$.

There are two main reasons why it is often impossible to include all important factors on the individual level into the analysis. Sometimes there are too many covariates to be considered in the model, in other cases the researcher do not know or is not able to measure all the relevant covariates. In both cases, there are two sources of variability in survival data: variability accounted for by measurable risk factors, which is thus theoretically predictable, and heterogeneity caused by unknown covariates, which is thus theoretically unpredictable, even if knowing all the relevant information. There are advantages to separate these two sources of variability since heterogeneity in contrast to variability can explain some "unexpected" results or can provide an alternative explanation of some results. Consider for example, non-proportional hazards or decreasing hazards when unexpected extra variability prevails.

In a proportional hazards model, neglect of a subset of the important covariates leads to biased estimates of both regression coefficients and the hazard rate. The reason for such bias lies in the fact that the time-dependent hazard rate results in changes in the composition of the study population over time with respect to the covariates.

If there are two groups of patients in a clinical trial where some individuals experience a higher risk of failure, then the remaining individuals at risk tend to form a more or less selected group

with a lower risk. An estimate of the individual hazard rate, without taking into account the unobserved frailty, would therefore underestimate the true hazard function and the extent of underestimation would increase as time progresses.

The univariate frailty model extends the Cox model such that the hazard of an individual depends in addition on an unobservable random variable Z , which acts multiplicatively on the baseline hazard function μ :

$$\mu(t,Z,X) = Z\mu_0(t)\exp(\beta^T X). \quad (1)$$

Again, $\mu_0(t)$ is the baseline hazard function, β the vector of regression coefficients, X is the vector of observed covariates. and Z now is the frailty variable. The frailty Z is a random variable varying over the population which lowers ($Z < 1$) or increases ($Z > 1$) the individual risk. Frailty corresponds to the notions liability or susceptibility in different settings (Falconer, 1967). The most important point here is that frailty is unobservable. The respective survival function S , describing the fraction of surviving individuals in the study population, is given by

$$S(t|Z, X) = \exp(-Z \int_0^t \mu_0(s) ds \exp(\beta^T X)) \quad (2)$$

$S(t|Z,X)$ may be interpreted as the fraction of individuals surviving the time t after begin of follow-up given the vector of observable covariates X and frailty Z . Note, that relations (1) and (2) describe the same model using different notions. Up to now, the model has been described at the level of individuals. However, this individual model is not observable. Consequently, it is necessary to consider the model at the population level. The survival function of the total population is the mean of the individual survival functions (2). It can be viewed as the survival function of a randomly drawn individual, and corresponds to that which is actually observed. It is important to note that the observed hazard function will not be similar to the individual hazard rate. What may be observed in the population is the net result for a number of individuals with

different Z . The population hazard rate may have a completely different shape than the individual hazard rate as shown in the following picture:

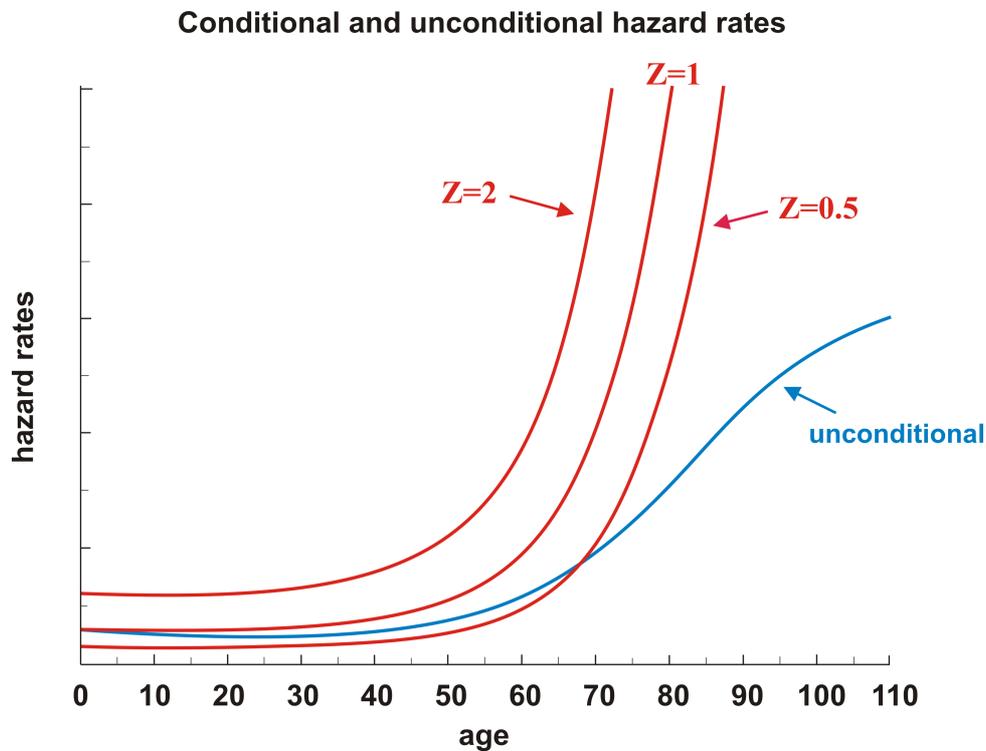


Figure 1: Conditional and unconditional hazard rates in a simulated data set of human mortality. The red lines denote the conditional (individual) hazard rates for individuals with frailty 0.5, 1 and 2, respectively. The blue line denotes the unconditional (population) hazard rate.

One important problem in the area of frailty models is the choice of the frailty distribution. The frailty distributions most often applied are the gamma distribution (Clayton 1978; Vaupel et al. 1979), the positive stable distribution (Hougaard 1986b), a three-parameter distribution (PVF) (Hougaard 1986a), the compound poisson distribution (Aalen 1988, 1992) and the log-normal distribution (McGilchrist and Aisbett, 1991).

Univariate frailty models are widely applied. A few examples which can be consulted for more details are listed here. Aalen and Tretli (1999) applied the compound Poisson distribution to

testicular cancer data already introduced above. The idea of the model is that a subgroup of men is particularly susceptible to testicular cancer, which results in selection over time.

Another example is the malignant melanoma data set including records of patients who had radical surgery for malignant melanoma (skin cancer) at the University Hospital of Odense in Denmark. Hougaard (2000) compared the traditional Cox regression model with a gamma frailty and PVF frailty model, respectively, to these data.

The third example deals with the time from insertion of a catheter into dialysis patients until it has to be removed due to infection. A subset of the complete data, including the first two infection times of 38 patients, was published by McGilchrist and Aisbett (1991). To account for heterogeneity within the data, Hougaard (2000) used a univariate gamma frailty model.

Henderson and Oman (1999) tried to quantify the bias which may occur in estimated covariate effects, and fitted marginal distributions when frailty effects are present in survival data but the latter are ignored in a misspecified proportional hazards analysis.

Congdon (1995) investigated the influence of different frailty distributions (gamma, inverse Gaussian, stable, binary) on total and cause-specific mortality from the London area (1988-1990).

Multivariate frailty models

A second important application of frailty models is in the field of multivariate survival data. Such kind of data occurs for example if lifetimes (or times of onset of a disease) of relatives (twins, parent-child) or recurrent events like infections in the same individual are considered. In such cases independence between the clustered survival times can not be assumed. Multivariate models are able to account for the presence of dependence between these event times. A commonly used and very general approach is to specify independence among observed data items conditional on a set of unobserved or latent variables (Hougaard, 2000). The dependence structure in the multivariate model arises from a latent variable in the conditional models for multiple observed survival times, for example let $S(t_1|Z, X_1)$ and $S(t_2|Z, X_2)$ be the conditional

survival functions of two related individuals with different vectors of observed covariates X_1 and X_2 , respectively, (see (2)). Averaging over an assumed distribution for the latent variables (e.g., using a gamma, log-normal, stable distribution) then induces a multivariate model for the observed data. In the case of paired observations, the two-dimensional survival function is of the form

$$S(t_1, t_2) = \int_0^{\infty} S(t_1|z, X_1)S(t_2|z, X_2)g(z)dz,$$

where g denotes the density of the frailty Z . In the case of twins, $S(t_1, t_2)$ denotes the fraction of twins pairs with twin 1 surviving t_1 and twin 2 surviving t_2 .

Frailty models for multivariate survival data are derived under conditional independence assumption by specifying latent variables that act multiplicatively on the baseline hazard.

The shared frailty model

The shared frailty model is relevant to event times of related individuals, similar organs and repeated measurements. Individuals in a cluster are assumed to share the same frailty Z , which is why this model is called shared frailty model. It was introduced by Clayton (1978) and extensively studied in Hougaard (2000). The survival times are assumed to be conditional independent with respect to the shared (common) frailty. For ease of presentation we will consider the case of groups with pairs of individuals (bivariate failure times, e.g. event times of twins or parent - child). Extensions to multivariate data are straightforward. Conditional on the frailty Z , the hazard function of an individual in a pair is of the form

$Z\mu_0(t)\exp(\beta^T X)$, where the value of Z is common to both individuals in the pair, and thus is the cause for dependence between survival times within pairs. Independence of the survival times within a pair corresponds to a degenerate frailty distribution ($Z=1, \sigma^2=0$). In all other cases with

$\sigma^2 > 0$ the dependence is positive by construction of the model. Conditional on Z , the bivariate survival function is given as

$$S(t_1, t_2 | Z) = S_1(t_1)^Z S_2(t_2)^Z.$$

In most applications it is assumed that the frailty distribution (i.e. the distribution of the random variable Z) is a gamma distribution with mean 1 and variance σ^2 . Averaging the conditional survival function produces under this assumption survival functions of the form

$$S(t_1, t_2) = (S_1(t_1)^{-\sigma^2} + S_2(t_2)^{-\sigma^2} - 1)^{1/\sigma^2}$$

Shared frailty explains correlation's between subjects within clusters. However, it does have some limitations. Firstly, it forces the unobserved factors to be the same within the cluster, which may not always reflect reality. For example, at times it may be inappropriate to assume that all partners in a cluster share all their unobserved risk factors. Secondly, the dependence between survival times within the cluster is based on marginal distributions of survival times. However, when covariates are present in a proportional hazards model with gamma distributed frailty the dependence parameter and the population heterogeneity are confounded (Clayton and Cuzick, 1985). This implies that the joint distribution can be identified from the marginal distributions (Hougaard, 1986a). Thirdly, in most cases, a one-dimensional frailty can only induce positive association within the cluster. However, there are some situations in which the survival times for subjects within the same cluster are negatively associated. For example, in the Stanford Heart Transplantation Study, generally the longer an individual must wait for an available heart, the shorter he or she is likely to survive after the transplantation. Therefore, the waiting time and the survival time afterwards may be negatively associated.

To avoid the above mentioned limitations of shared frailty models correlated frailty models were developed.

The correlated frailty model

Originally, correlated frailty models were developed for the analysis of bivariate failure time data, in which two associated random variables are used to characterize the frailty effect for each pair. For example, one random variable is assigned for partner 1 and one for partner 2 so that they would no longer be constrained to have a common frailty. These two variables are associated and have a joint distribution. Knowing one of them does not necessarily imply knowing the other. There is no more a restriction on the type of correlation. These two variables can also be negatively associated, which would induce a negative association between survival times. Assuming gamma distributed frailties, Yashin and Iachine (1995) used the correlated gamma-frailty model resulting in a bivariate survival distribution of the form

$$S(t_1, t_2) = \frac{S_1(t_1)^{1-\rho} S_2(t_2)^{1-\rho}}{(S_1(t_1)^{-\sigma^2} + S_2(t_2)^{-\sigma^2} - 1)^{\rho/\sigma^2}}.$$

Examples of the use of multivariate frailty models are various and emphasize the importance of this family of statistical models for survival data.

- a shared log-normal frailty model for the catheter infection data mentioned above used by McGilchrist and Aisbett (1991),
- a shared frailty model with gamma and log-normal distributed frailty, applied to the recurrence of breast cancer by dos Santos et al. (1995),
- a shared positive stable frailty model, applied by Manatunga and Oakes (1999) to the data from the Diabetic Retinopathy Study, which examined the effectiveness of laser photo-coagulation in delaying the onset of blindness in patients with diabetic retinopathy. Positive stable frailty allows for proportional hazards both in the marginal and the conditional model.
- a study of Andersen et al. (1999), who tested for centre effects in multi-centre survival studies by means of a frailty model with unspecified frailty distribution,
- a correlated gamma-frailty model, applied by Pickles et al. (1994) to age of onset of puberty and antisocial behavior in British twins,

- a correlated gamma-frailty model by Yashin and Iachine (1995) and Yashin et al. (1995) to analyze mortality in Danish twins
- a correlated gamma-frailty model by Wienke et al. (2001) and Zdravkovic et al. (2002) to analyze genetic factors involved in mortality due to coronary heart disease in Danish and Swedish twins, respectively,
- an extension of the correlated gamma-frailty model by Wienke et. al (2002a) used to model death due to coronary heart disease in Danish twins,
- different versions of the correlated gamma-frailty model applied by Zahl (1997) to cause-specific cancer mortality in Norway to model the excess hazard.

Software

Stata 7 (procedure `st streg`) allows to explore univariate models with gamma and inverse Gaussian distributed frailty. `aML 2` supports log-normal frailty models in univariate analysis. `WinBugs` is designed for analysis of shared frailty models with different frailty distribution, using Markov Chain Monte Carlo methods. In the internet, there are several SAS, GAUSS, S-Plus and R routines available dealing with different frailty models.

References

1. Aalen, O.O. (1988) Heterogeneity in Survival Analysis. *Statistics in Medicine* 7, 1121 - 1137
2. Aalen, O.O. (1992) Modelling Heterogeneity in Survival Analysis by the Compound Poisson Distribution. *Annals of Applied Probability* 4 (2), 951 - 972
3. Aalen, O.O., Tretli, S. (1999) Analysing incidence of testis cancer by means of a frailty model. *Cancer Causes and Control* 10, 285 - 292
4. Andersen, P.K., Klein, J.P., Zhang, M.-J. (1999) Testing for Centre Effects in Multi-Centre Survival Studies: A Monte Carlo Comparison of Fixed and Random Effects Tests. *Statistics in Medicine* 18, 1489 - 1500

5. Clayton, D.G. (1978) A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65, 141-151
6. Clayton, D., Cuzick, J. (1985) The semi-parametric Pareto model for regression analysis of survival times. *Proceedings of the Centenary Session of the International Statistical Institute, Amsterdam*
7. Congdon, P. (1995) Modelling Frailty in Area Mortality. *Statistics in Medicine* 14, 1859-74
8. Cox, D.R. (1972) Regression Models and Life-Tables. *Journal of the Royal Statistical Society B* 34, 187-220
9. dos Santos, D.M., Davies, R.B., Francis, B. (1995) Nonparametric hazard versus nonparametric frailty distribution in modelling recurrence of breast cancer. *Journal of Statistical Planning and Inference* 47, 111-127
10. Falconer, D.S. (1967) The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. *Annals of Human Genetics* 31, 1 - 20
11. Greenwood, M., Yule, G.U. (1920) An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society* 83, 255-79
12. Henderson, R., Oman, P. (1999) Effect of frailty on marginal regression estimates in survival analysis. *Journal of the Royal Statistical Society B* 61, 367 - 379
13. Hougaard, P. (1986a) Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, 671 - 678
14. Hougaard, P. (1986b) A class of multivariate failure time distributions. *Biometrika* 73, 671-78
15. Hougaard, P. (1991) Modelling Heterogeneity in Survival Data. *Journal of Applied Probability* 28, 695 - 701
16. Hougaard, P. (2000) *Analysis of multivariate survival data*. Springer, New York
17. Manton, Stallard (1981) Methods for evaluating the heterogeneity of aging processes in human populations using vital statistics data: explaining the black/white mortality crossover by a model of mortality selection. *Human Biology* 53, 47 - 67
18. Manatunga, A.K., Oakes, D. (1999) Parametric Analysis of Matched Pair Survival Data. *Lifetime Data Analysis* 5, 371 - 387

19. McGilchrist, C.A., Aisbett, C.W. (1991) Regression with Frailty in Survival Analysis. *Biometrics* 47, 461-466
20. Pickles, A., Crouchley, R., Simonoff, E., Eaves, L., Meyer, J., Rutter, M., Hewitt, J., Silberg, J. (1994) Survival Models for Developmental Genetic Data: Age of Onset of Puberty and Antisocial Behavior in Twins. *Genetic Epidemiology* 11, 155 - 170
21. Vaupel, J.W., Manton, K.G., Stallard, E. (1979) The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. *Demography* 16, 439 - 454
22. Wienke, A., Holm, N., Skytthe, A., Yashin, A.I. (2001) The heritability of mortality due to heart diseases: a correlated frailty model applied to Danish twins. *Twin Research* 4, 266 - 274
23. Wienke, A., Christensen, K., Skytthe, A., Yashin, A.I. (2002a) Genetic analysis of cause of death in a mixture model with bivariate lifetime data. *Statistical Modelling* 2, 89 - 102
24. Wienke, A., Herskind, A.M., Christensen, K., Skytthe, A., Yashin, A.I. (2002b) The influence of smoking and BMI on heritability in susceptibility to coronary heart disease. Max Planck Institut für demografische Forschung, Working Paper 2002-03 (www.demogr.mpg.de)
25. Wienke, A., Arbeev, K., Locatelli, I., Yashin, A.I. (2003a) A simulation study of different correlated frailty models and estimation strategies. Max Planck Institut für demografische Forschung, Working Paper 2003-018 (www.demogr.mpg.de)
26. Wienke, A., Lichtenstein, P., Yashin, A.I. (2003b) A bivariate frailty model with cure fraction for modeling familial correlations in diseases. (*Biometrics*, forthcoming)
27. Yashin, A.I., Iachine, I.A. (1995) Genetic Analysis of Durations: Correlated Frailty Model Applied to Survival of Danish Twins. *Genetic Epidemiology* 12, 529 - 538
28. Yashin, A.I., Vaupel, J.W., Iachine, I.A. (1995) Correlated Individual Frailty: An Advantageous Approach to Survival Analysis of Bivariate data. *Mathematical Population Studies* 5, 145 - 159
29. Zahl, P.H. (1997) Frailty modelling for the excess hazard. *Statistics in Medicine* 16, 1573-85
30. Zdravkovic, S., Wienke, A., Pedersen, N.L., Marenberg, M.E., Yashin, A.I., de Faire, U. (2002) Heritability of death from coronary heart disease: a 36-year follow-up of 20,966 Swedish twins. *Journal of Internal Medicine* 252, 247 - 54