



UNIVERSITETET  
I OSLO

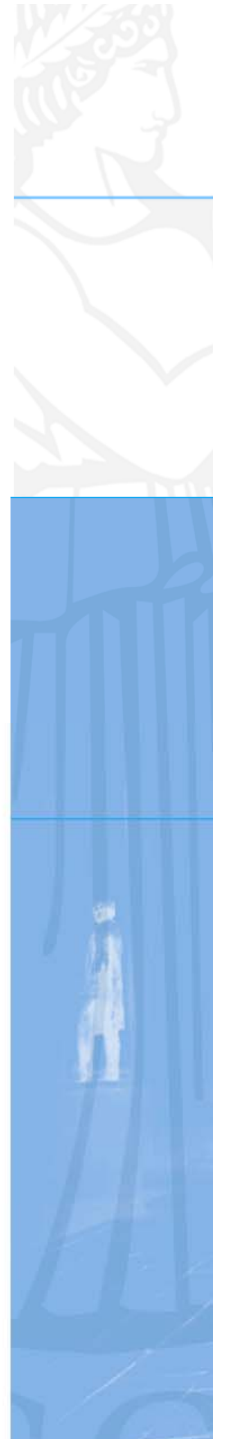
# A Frailty Analysis Of Two-Generation Melanoma Data From The Swedish Multi-Generation Register

Tron Anders Moger

Institute of Health Management and Health  
Economics

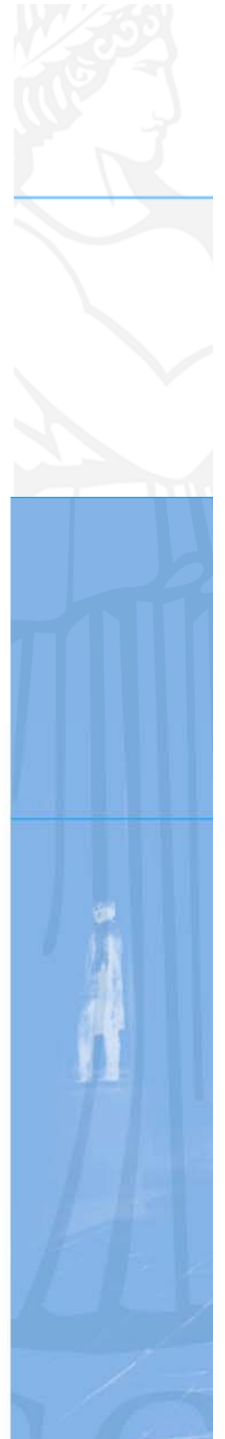
IMB/Dept. of Biostatistics

University of Oslo



# Project group

- Joint work with:
- Marion Haugen (IMB/Department of biostatistics, University of Oslo, Norway)
- Håkon K. Gjessing (Norwegian Institute of Public Health, Norway)
- Ørnulf Borgan (Department of statistics, University of Oslo, Norway)
- Ben Yip (University of Hong Kong, Hong Kong)
- Yudi Pawitan (Dep. of medical epidemiology and biostatistics, Karolinska, Sweden)



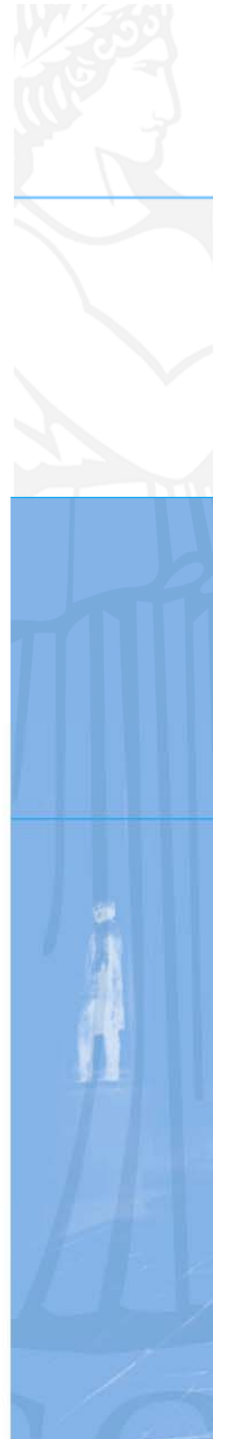
# Aim:

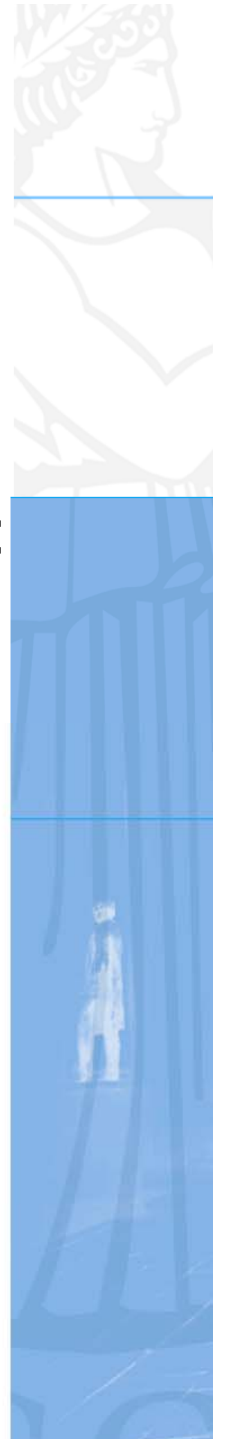
- Develop new hierarchical frailty models for family data based on Moger and Aalen (2005)
- Alternative to e.g. additive frailty models or multivariate lognormal frailty models
- Use a melanoma data set from the Swedish multi-generation register to illustrate the potential use of these models



# Contents

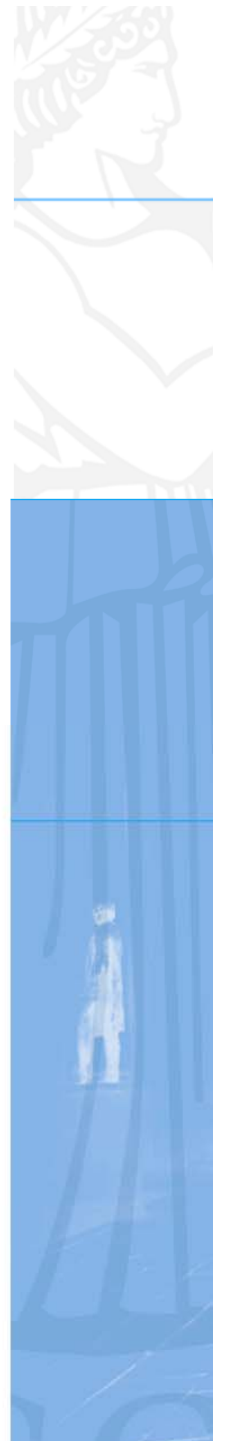
- Introduction to frailty models
- Presentation of the model
- Presentation of the melanoma data
- Results





# Introduction to frailty models

- Multiplicative frailty model (see also Hougaard, 2000):  
Individual hazard  $= Z^* \lambda(t)$
- $Z$  is a frailty variable, measures unobserved heterogeneity,  $\lambda(t)$  is a baseline hazard
- $Z$  has correlated values within families
- High value of  $Z$  implies high risk of disease
- Survival function  $S(t) = \int \exp(-z\Lambda(t))g(z)dz = L_Z(\Lambda(t))$
- Covariates (observed heterogeneity) often included in the model



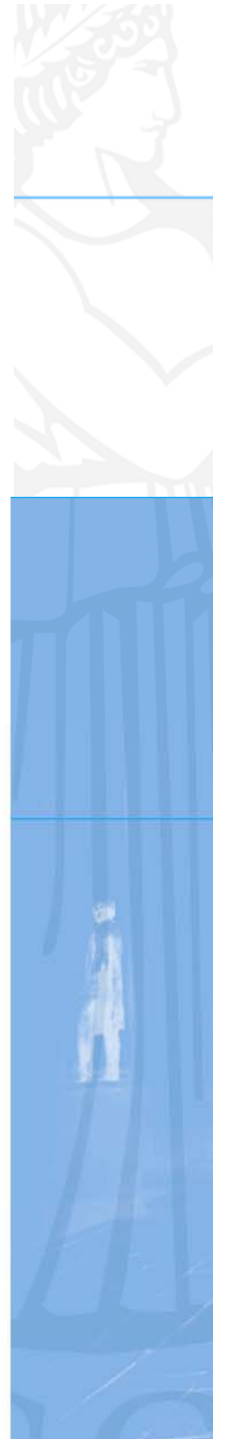
## Frailty models; family data

- Dependence structure might be complex, e.g. for two-generation families:
  - Individual environment
  - Genetics: Parents independent, parents and children have correlation  $\frac{1}{2}$ , within-children correlation is also  $\frac{1}{2}$
  - Common environment (For melanoma: sun exposure?)
- Need a frailty model with several levels of dependence



# Common frailty distributions

- Gamma, stable, inverse Gaussian and PVF distributions
- They all have a Laplace transform which can be written as  $L_z(s) = \exp(-\rho\Psi(s))$ 
  - $\rho$  is a scale parameter
  - $\psi(s)$  is the Laplace exponent



# Generalization of the frailty distribution

- Let the frailty on each level be given by  $Z_1, Z_2, Z_3$  etc.
- Write Laplace transform for level 1 as:

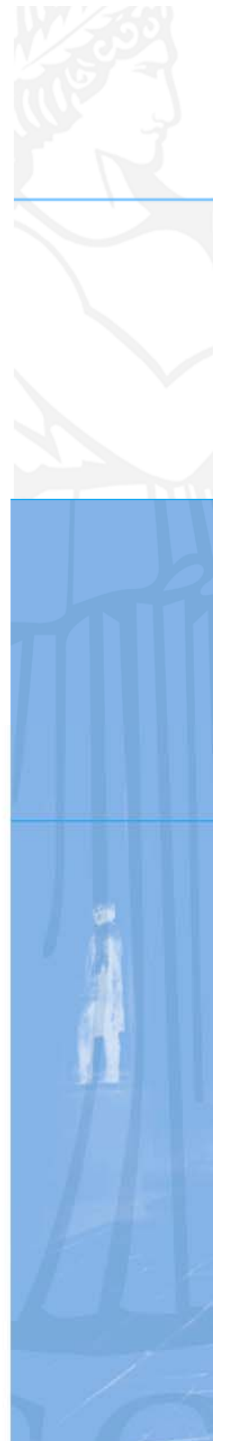
$$L_{Z_1}(s) = \exp(-z_2 \Psi_1(s))$$

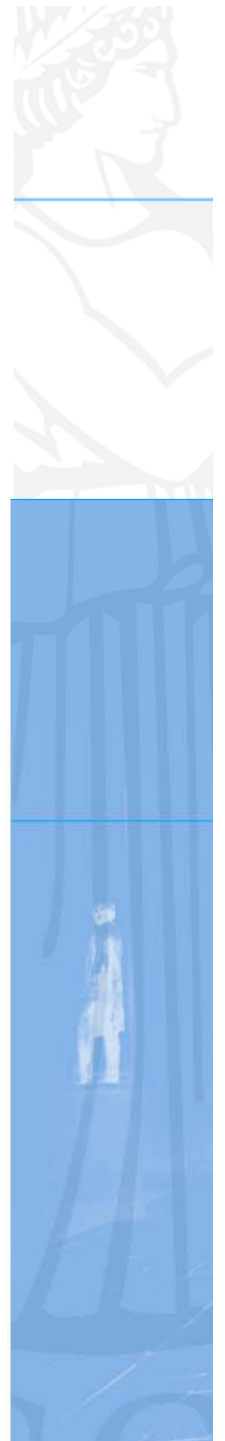
- Randomize the parameter  $z_2$  by another frailty variable  $Z_2$  with Laplace transform  $L_{Z_2}(s) = \exp(-z_3 \Psi_2(s))$
- Given the frailties on higher levels,  $z_1, z_2$  etc. are fixed, shared by some and independent for others
- For a three level model, the Laplace transform for each individual becomes  $L(s) = \exp(-z_4 \Psi_3(\Psi_2(\Psi_1(s))))$



# Construction of a model with environmental and genetic frailties

- $Z_1$  is independent for all, with  $L_{Z_1}(s) = \exp(-z_2 \Psi_1(s))$
- Apply an additive genetic component to the parameter  $z_2$ , which looks like the following for a mother (M), a father (F) and two children ( $C_1$  and  $C_2$ ):
  - $Z_{2M} = M_1 + M_2 + M_3 + M_4$
  - $Z_{2F} = F_1 + F_2 + F_3 + F_4$
  - $Z_{2C1} = M_2 + M_3 + F_2 + F_3$
  - $Z_{2C2} = M_3 + M_4 + F_3 + F_4$
- All  $M_i, F_i$  are iid with Laplace transform
$$L(s) = \exp(-z_3 \Psi_2(s))$$





## Laplace transform:

- By conditioning on the frailties and integrating them out, one gets the joint Laplace transform (and survival function+density):

$$\begin{aligned} L(m, f, c_1, c_2) = & \exp(-z_3[\Psi_2(\Psi_1(m)) + \Psi_2(\Psi_1(f)) \\ & + \Psi_2(\Psi_1(m) + \Psi_1(c_1)) + \Psi_2(\Psi_1(f) + \Psi_1(c_1)) \\ & + \Psi_2(\Psi_1(m) + \Psi_1(c_2)) + \Psi_2(\Psi_1(f) + \Psi_1(c_2)) \\ & + \Psi_2(\Psi_1(m) + \Psi_1(c_1) + \Psi_1(c_2)) \\ & + \Psi_2(\Psi_1(f) + \Psi_1(c_1) + \Psi_1(c_2))]) \end{aligned}$$



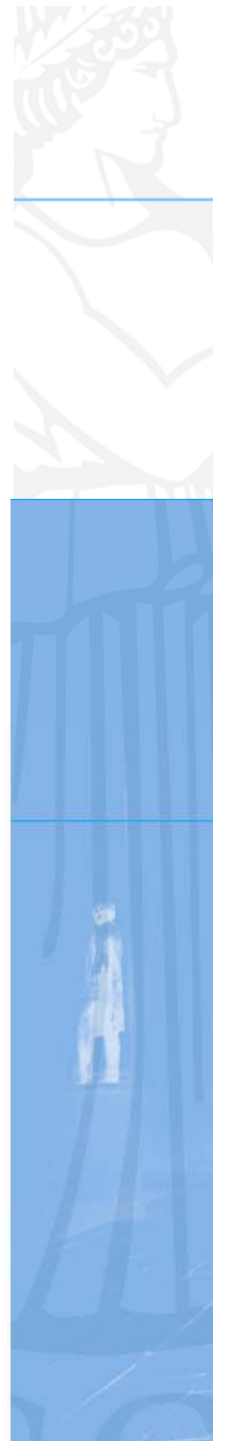
# Common environment

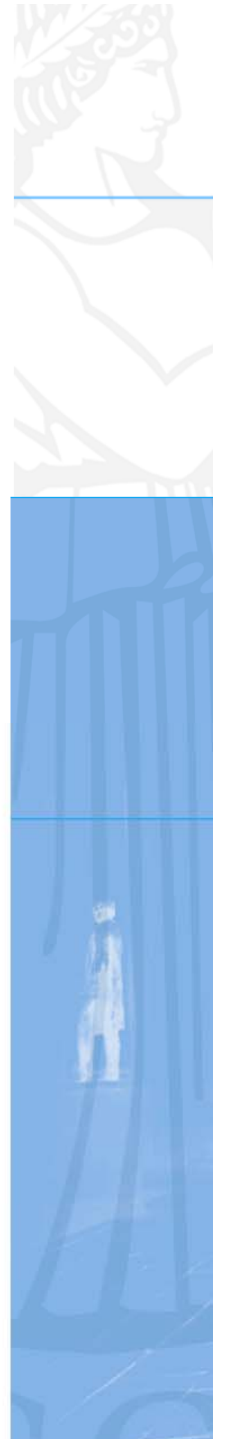
- Can apply a common environmental term to  $z_3$ , with Laplace transform

$$L_{z_3}(s) = \exp(-z_4 \Psi_3(s))$$

- Get joint Laplace transform

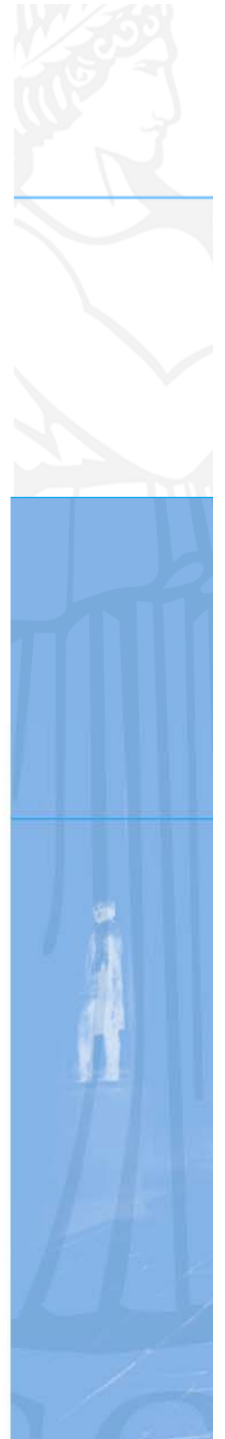
$$\begin{aligned} L(m, f, c_1, c_2) = & \exp(-z_4 [\Psi_3(\Psi_2(\Psi_1(m)) + \Psi_2(\Psi_1(f))) \\ & + \Psi_2(\Psi_1(m) + \Psi_1(c_1)) + \Psi_2(\Psi_1(f) + \Psi_1(c_1)) \\ & + \Psi_2(\Psi_1(m) + \Psi_1(c_2)) + \Psi_2(\Psi_1(f) + \Psi_1(c_2)) \\ & + \Psi_2(\Psi_1(m) + \Psi_1(c_1) + \Psi_1(c_2)) \\ & + \Psi_2(\Psi_1(f) + \Psi_1(c_1) + \Psi_1(c_2))]) \end{aligned}$$





# Important properties of the model

- All individuals get same distributions marginally
- Also, they have the same expected value and variance
- Correlation structure is "realistic"
- Let  $Y$  be the total frailty for each individual
- Can also show that, under certain conditions,  
 $E(Y)=1$  and  $\text{Var}(Y)=\text{Var}(Z_1)+\text{Var}(Z_2)+\text{Var}(Z_3)$



# The heritability coefficient $h^2$

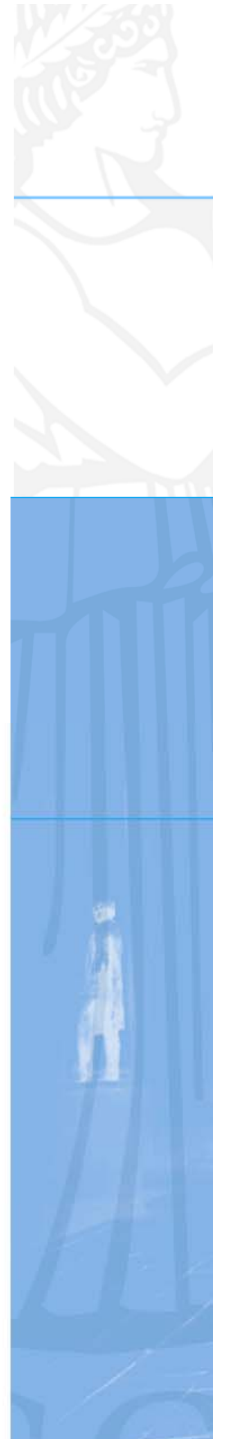
- Heritability coefficient: Proportion of total frailty variance due to genetics
- Easy to find the heritability coefficient:
  - Let  $\sigma^2_1$ ,  $\sigma^2_2$ ,  $\sigma^2_3$  be the variance on the level for individual environment, genetics and common environment, respectively
  - Then  $h^2$  is  $\sigma^2_2/(\sigma^2_1+\sigma^2_2+\sigma^2_3)$



# Frailty relative risk as a measure of dependence

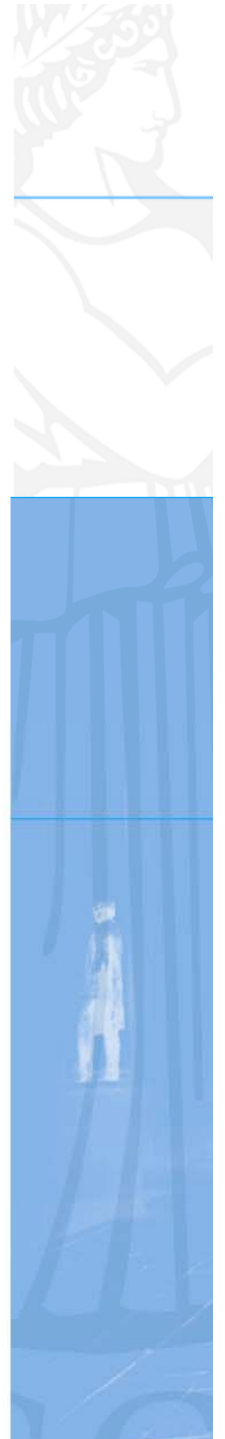
$$RR = \frac{P(\text{Ind. gets event within age } t \mid \text{Relative gets event within age } t)}{P(\text{Ind. gets event within age } t \mid \text{Relative no event up to age } t)}$$

- The probabilities can be calculated by means of bivariate and univariate Laplace transforms, using the fact that individuals are independent given  $Z_2$  and  $Z_3$
- Gives a measure of dependence due to common frailty also after adjusting for covariates



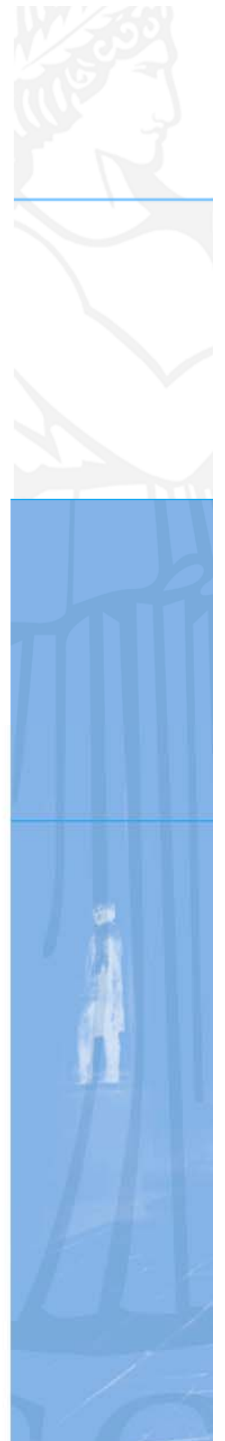
# Alternatives to our model

- Could have decomposed total frailty  $Y$  into additive terms for individual and common environment, and genetics (as e.g in Korsgaard and Andersen, 1998)
- Interesting to see how this model compares to a purely additive model
- Will get back to a comparison to additive frailty models



## Melanoma data:

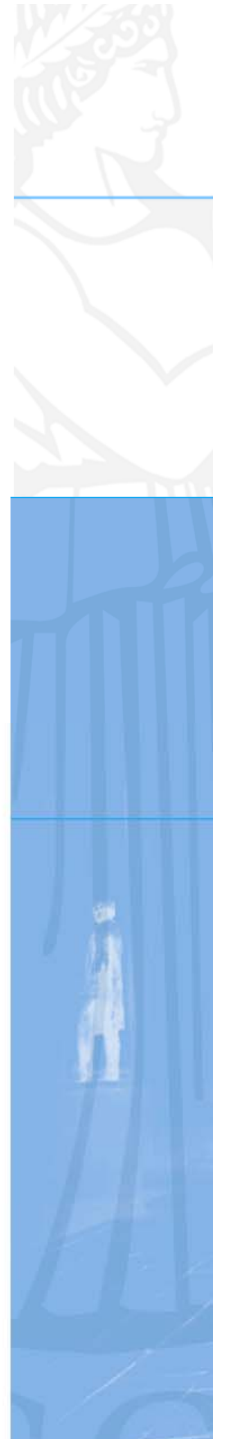
- Will analyze melanoma data from the Swedish Multi-Generation register
- Includes 10.3 million individuals organized in 3 million families
- Too much data; need to analyze case-cohort sample (Moger et al., 2008)
- Would like to quantify the dependence
- And look at how much each frailty component contributes to the total frailty variation



## Details of the sample:

- Around 400000 individuals in the sample (out of 10.3 million)
- Only look at parents + two oldest children
- Covariates are birth cohort and gender
- Distribution of cases:

| Affected: | Parents: |         |         |       |
|-----------|----------|---------|---------|-------|
| Children: | None:    | Mother: | Father: | Both: |
| None      | 126196   | 11125   | 11016   | 73    |
| 1st child | 6047     | 90      | 91      | 3     |
| 2nd child | 2893     | 39      | 36      | 0     |
| Both      | 37       | 3       | 1       | 0     |



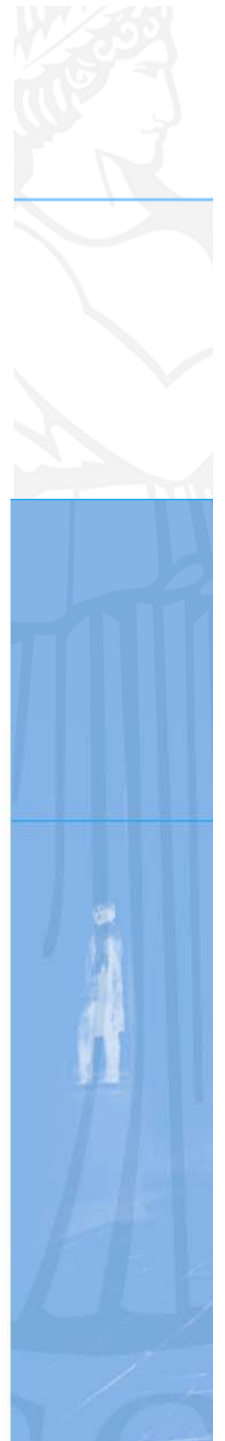
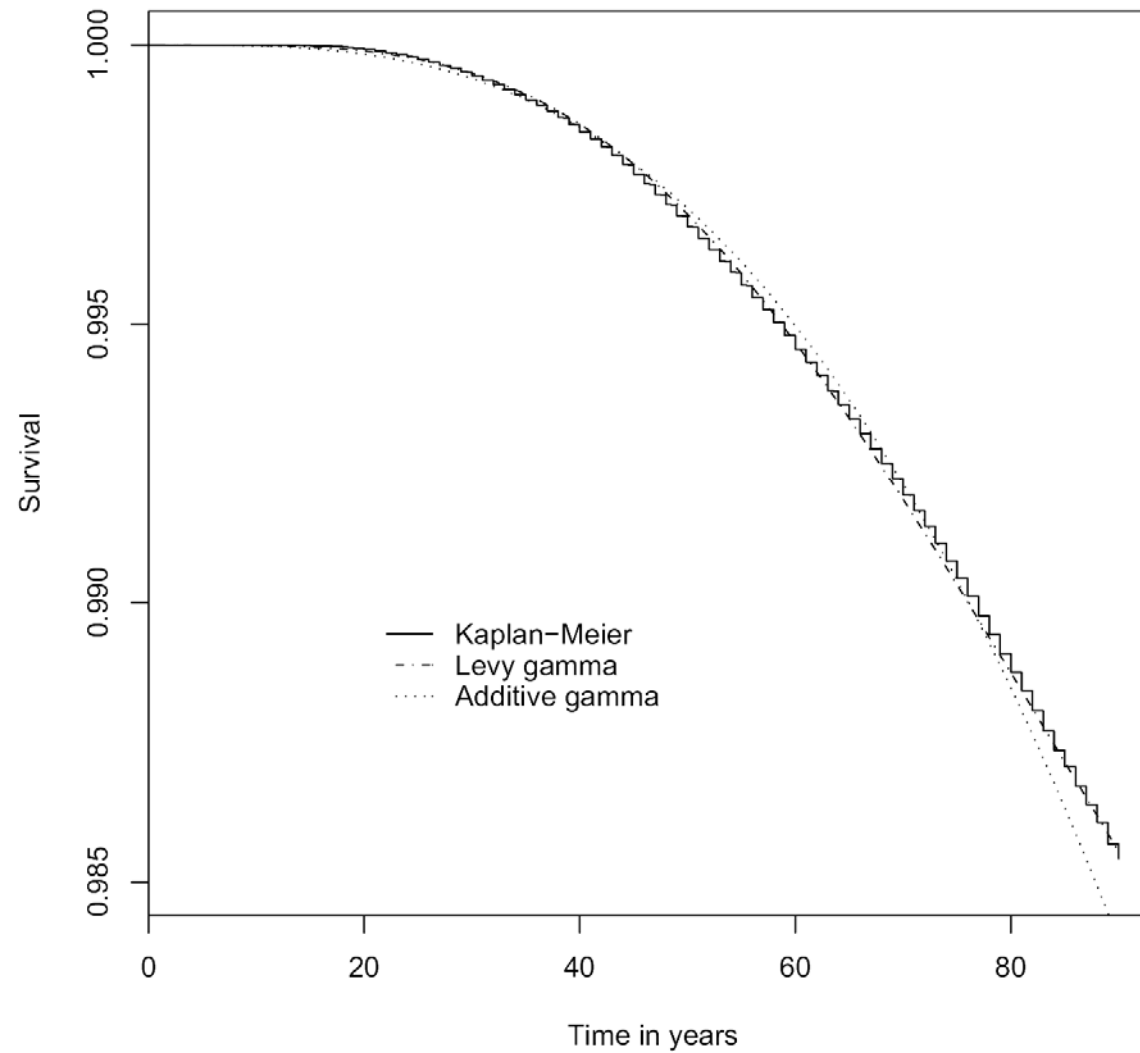
# Analysis

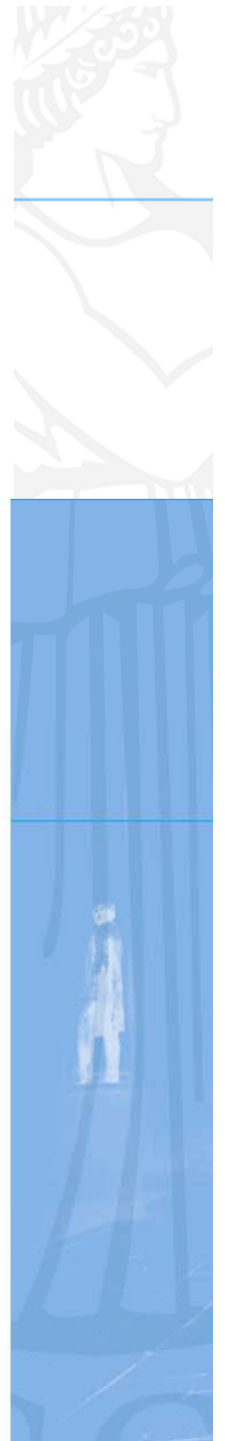
- Put the covariates in a Cox-regression term in the baseline hazard  $\lambda(t)$
- Baseline hazard is Weibull
- Use gamma distributions for all frailty components
- Have tried other distributions also, appear to give similar results
- Gives 5 parameters in model+regression coefficients (One for each frailty level+two Weibull parameters)
- Have also tried an additive gamma model with 5 parameters and same dependence structure



# Model fit appears to be good:

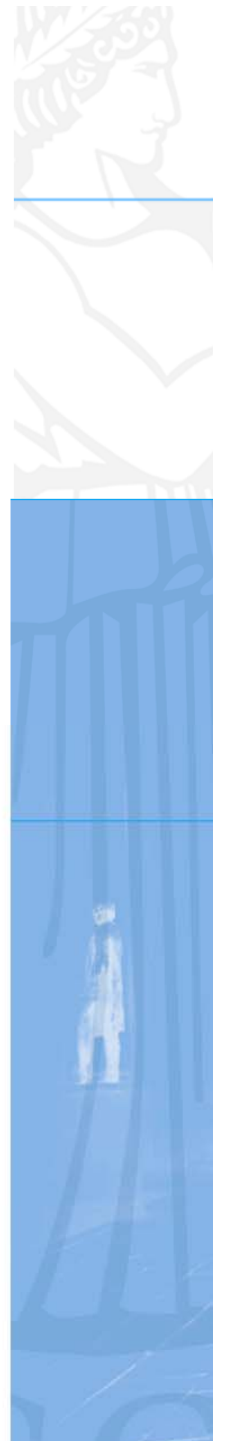
All family members





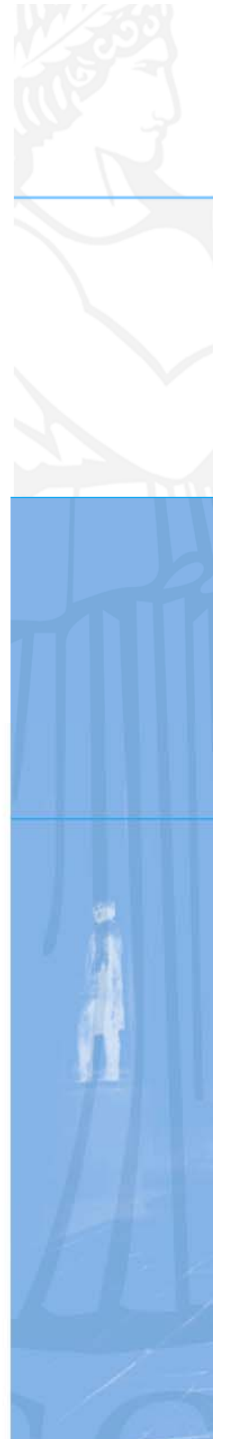
## Results without covariates:

- Genetic variance:  $\sigma^2_2=2.00$  (SE=0.49)
- Common envir. variance:  $\sigma^2_3=0.66$  (SE=0.19)
- Individual envir. variance:  $\sigma^2_1=93$  (SE=4.2)
- See that the genetics does not contribute much to the total frailty variance,  $h^2$  is around 2%
- Can be compared to an estimate from the Melanoma Genetics Consortium (Kefford et al., 1999), saying that the proportion of all melanoma attributable to susceptibility genes is at most 2%



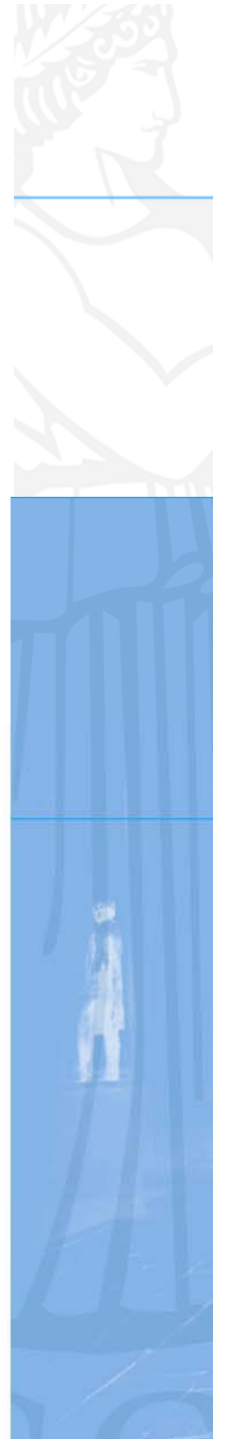
## Results with covariates

- Relative risk for birth cohort per 10 yrs: 1.49 (95%CI: 1.48-1.51)
- Relative risk for males vs females: 0.97 (0.94-0.99)
- Genetic variance: 3.1 (SE=0.43)
- Common environmental variance: 0.23 (SE=0.14)
- Individual environmental variance: 0.001 (SE=0.06)
- Genetics now account for 93.0% of the total frailty variance
- Similarly, common environment accounts for 6.9% and individual environment for 0.1%



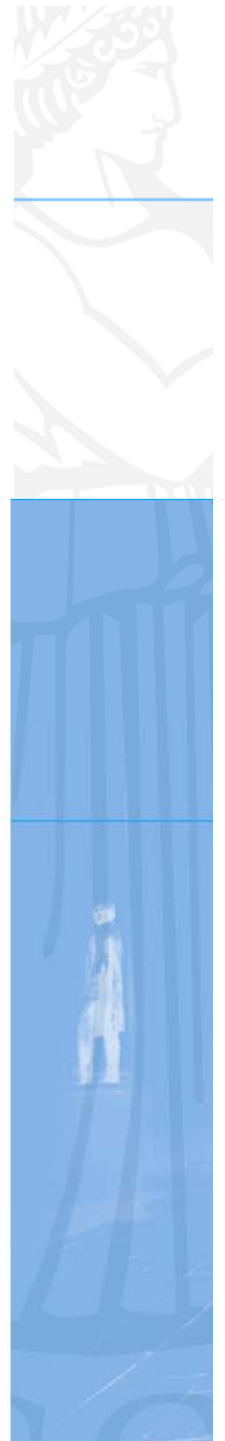
## Frailty relative risk:

- If your parent, child or sibling has melanoma within 90 yrs of age, you have a 2.71 times higher risk of developing it yourself
- This is within the range of other studies published (e.g. Hemminki et al, 2003, gets 2.40 for parent child pairs and 2.98 for sibling pairs)
- If your partner has melanoma, you have a 1.23 times higher risk of becoming affected yourself
- For obvious reasons, probably not many have estimated this



## Conclusion

- Have presented a new frailty model with a hierarchical dependence structure
- The model has the same nice properties as additive frailty models
- Appears to give a better fit than an additive frailty model to the melanoma data
- Joint frailty analyses of family data are not that common due to the complexity of the models, hence the application should be interesting



# References

- Hemminki, K., Zheng, H. and Czene, K. Familial and attributable risks in cutaneous melanoma: Effects of proband and age. *Melanoma Res*, **11**: 133-40 (2003).
- Hougaard, P., *Analysis of Multivariate Survival Data*. New York: Springer Verlag (2000).
- Kefford et al. Counseling and DNA testing for individuals perceived to be genetically predisposed to melanoma: a consensus statement of the Melanoma Genetics Consortium. *J Clin Oncol*, **17**: 3245-51 (1999).
- Korsgaard I. R. and Andersen A. H., The additive genetic frailty model. *Scand J Stat*, **25**, 255-69 (1998).
- Moger T. A., Aalen O. O., A distribution for multivariate frailty based on the compound Poisson distribution with random scale, *Lifetime Data Analysis*, **11**, 41-59 (2005).
- Moger T. A., Pawitan, Y., Borgan, Ø. Case-cohort methods for survival data on families from routine registers, *Statistics in Medicine* **27**, 1062-1074 (2008).