Session 4. Clustered data, frailties and score functions

Content:

- Martingale residuals
- Approximate likelihoods
- Centre effects
- Goodness-of-fit
- Pedigrees

Martingale residuals

First law of survival analysis

$$\mathrm{E}[d - H(T)] = 0$$ ($d \& H(T)$ is called the martingale residual)

d: event indicator (d=1 if dead, d=0 if censored)

H: cumulative hazard

T: censoring or survival time

Proof: an exercise

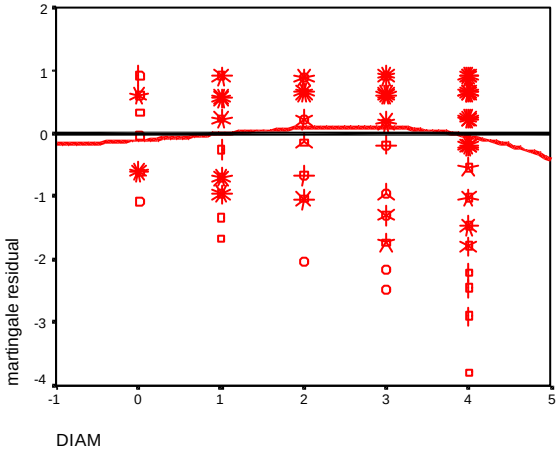
Second law

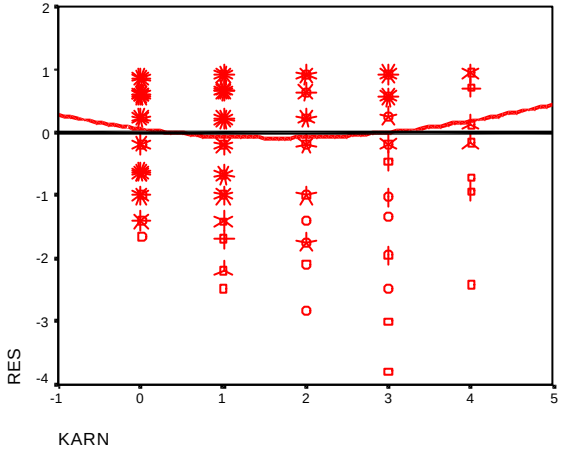
$$\mathrm{E}[d - H(T)]^2 = \mathrm{E}[H(T)]$$

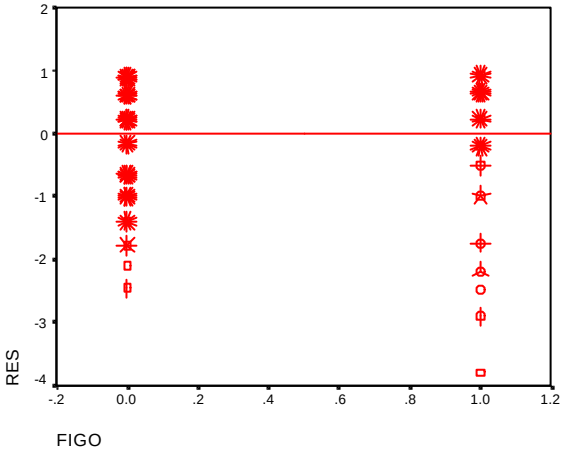
First application in graphical goodness of fit check.

Back to the ovarian cancer example.

Martingale residual after fitting a simple linear model.







Second application: simple validity check

Group of n individuals, observed (d_i, T_i)

(Historic) model states: cumulative hazard of individual i $H_i(t)$

Test on overall shift in survival based on

$$\frac{\sum (d_i - H_i(T_i))}{\sqrt{\sum H_i(T_i)}} \sim N(0,1) \quad \text{or} \quad \frac{[\sum (d_i - H_i(T_i))]^2}{\sum H_i(T_i)} \sim ?^2_{[1]}$$

NB. H, not 1-S !!!

Aproximate likelihood

Question:

What do the martingale residuals tell us, if the model does not fit.

Answer:

Ask Yusuf-Peto about their trick in meta-analysis

Model:

Group has relative risk $RR = e^{?}$ with respect historic model.

For each individual $S_i(t) = \exp(-e^{?} H_i(t))$

Log-likelihood of all data in the group w.r.t. ?

$$l(?) \,'\, \&e^{?} \,'\, H_i(T_i)\%? \,'\, d_i$$

Same as the Poisson-likelihood with $E \,'\, ' H_i(T_i)$ and $O \,'\, ' (d_i)$

Taylor-approximation around zero

$$l(?) \,'\, l(0)\%? \,'\, (d_i \& H_i(T_i))\&½?^2 \,'\, H_i(T_i)$$

Same likelihood is generated by

$$\hat{?} \,'\, \frac{' (D_i \& H_i(T_i))}{' H_i(T_i)} \sim N(?, \frac{1}{' H_i(T_i)})$$

From this approximation it is easy to derive the score-test for the (fixed) effect of a G-categorical covariate

$$\sum_{i=1}^{G} \frac{\left(\sum_{j} (d_{ij} - \hat{H}(T_{ij}))\right)^2}{\sum_{j} \hat{H}(T_{ij})} \sim \chi^2_{[G-1]}$$

Using

$$\hat{H}(t) = \sum_{t_i \le t} \frac{d_i}{r_i}$$, the pooled Breslow-estimator

leads to a version of the log-rank test.

For testing group-effects after correction for covariates, $\hat{H}$ is the estimated cumlative hazard in the model without group effects

Random effect versus fixed effects

$?_1,...,?_G$ true effects in the G groups/categories

< Fixed-effect approach tests $?_1 = ?_2 = ... = ?_G$ by test given above

(or LRT or Wald-test)

< Random-effect approach assumes $? \sim N(\mu, t^2)$ and test

$t^2 = 0$ versus $t^2 > 0$ by score-test or LRT-test

Score-test voor random effect is based on

$$\sum_{i=1}^{G}\Big[\sum_{j}(d_{ij}-H_{ij}(T_{ij}))^2-\sum_{j}(H_{ij}(T_{ij})\Big]$$

The cut-off value is harder to find

(Commenges and Andersen, 1995, Verweij and Van Houwelingen, 1998)

Random-effect test put little weight on small “centres”

(Andersen, Klein, Zhang, 1999)

Random-effect test has some local optimality property.

Multi-level modeling requires random effects

Random-effect analysis by means of the approximate likelihood

$$\hat{?} \sim \frac{\sum(D_i \& H_i(T_i))}{\sum H_i(T_i)} \sim N(?, \frac{1}{\sum H_i(T_i)})$$

is rather straight-forward, but a little suspect .

Demonstrated on Eurotransplant data

(Jaqueline Smits, ISCB2000)

QUALITY ASSESSMENT OF HEART TRANSPLANT CENTERS IN EUROTRANSPLANT

Jacqueline MA Smits
Eurotransplant International Foundation
Leiden, The Netherlands
JSMITS@eurotransplant.nl

Hans C van Houwelingen
LUMC
Leiden, The Netherlands

ISCB 0C-page1

Contents

- About the Eurotransplant data
- Aim of the study
- A score statistic approach
- Representation and Interpretation
- Conclusions and next steps

ISCB 00-page2

Data

- heart transplants performed between Jan.'97 and Dec.'97 (N=723) in the Eurotransplant countries
- 44 transplant centers, volume between 1 and 73 operations
- time till event data

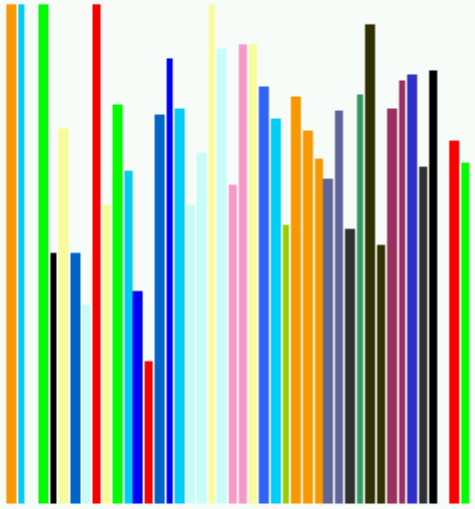
General purpose of the study

Quality assessment of heart transplant centers

Current situation

Observed 1 year patient survival rates

%



Future situation
as demanded by
national transplantation laws

- equality of access
- equality of success

ISCB 0C-page 5

Equality of success

=

Absence of center effect

ISCB00-page 6

Aim of the study

Study of center effects

- **corrected for covariates at patient level**
- **relating to covariates at center level**

Setting

Data: d_{ij} T_{ij}
Model: Cox PH

$$\ln(h_{ij}(t)) = \ln(h_0) + \theta_i + X_{ij}\beta$$

i: center
j: patient

Approach:

1. Ignore θ_i 's and fit β 's
2. Estimate θ_i 's

Center effect

Definition

Estimate

if $\theta = \ln(RR)$ and $\theta_i \approx 0$

then $\hat{\theta}_i = \dfrac{\sum_j d_{ij} - \sum_j H_{ij}(T_{ij})}{\sum_j H_{ij}(T_{ij})}$

ISCB00-page 9

Covariates

- Categorical
  - gender patient (2)
  - gender donor (2)
  - primary disease (2)
  - urgency class (3)
  - VAD (2)
  - respirator (2)
  - inotropics (2)
- Continuous
  - age patient (years)
  - age donor (years)
  - donor:recipient weight match
  - cold ischemia time (hours)
  - serum creatinine (mg/dl)

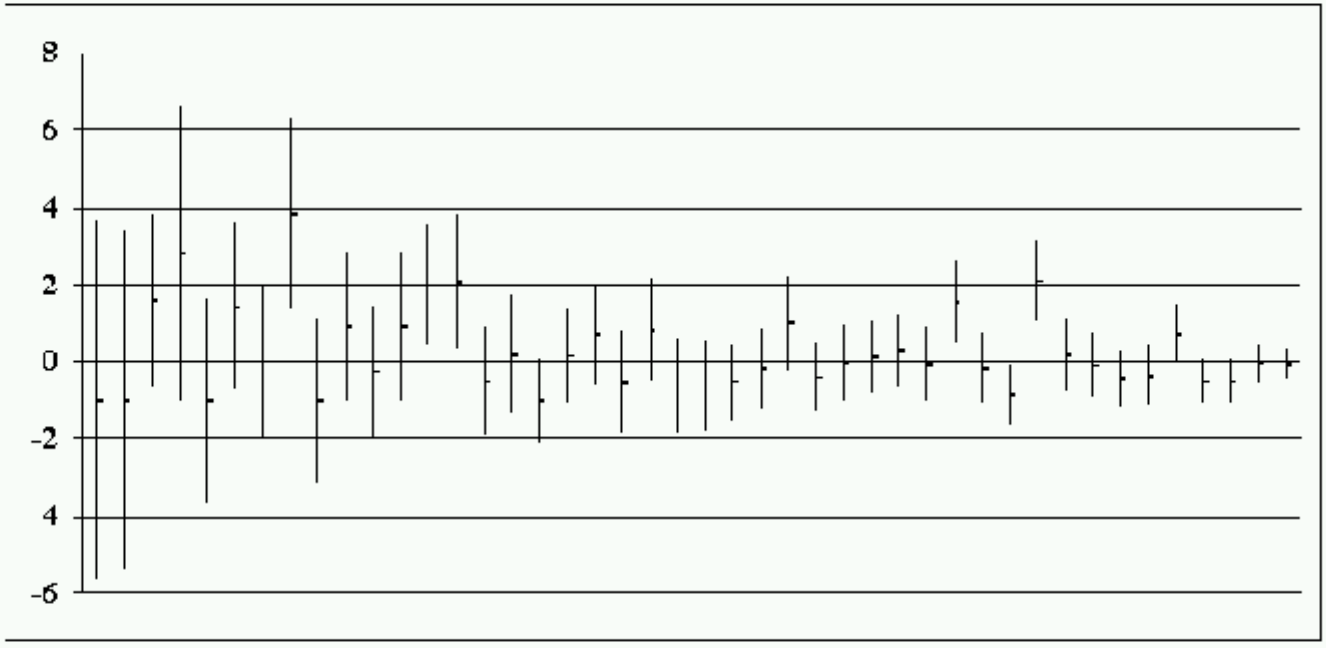
ISCB00-page ´0

Crude center effect

$$\hat{\theta}_i \sim N(\theta_i, \frac{1}{\sum_j H_{ij}})$$

$$95\% CL = (\hat{\theta}_i \pm 1.96 s_i)$$

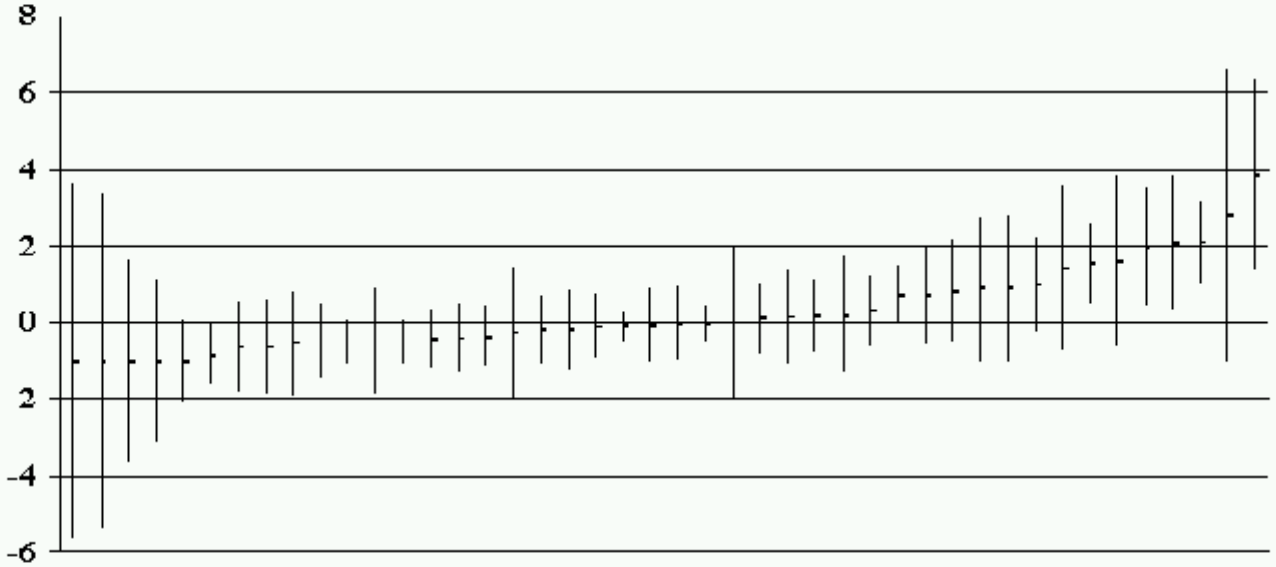
Crude center effects



$$\hat{\theta}_i \pm 1.96 s_i$$

ISCB00-page 12

Crude center effects



$$\hat{\theta}_i \pm 1.96 s_i$$

ISCB00-page 12

Problem

too much noise within centers

Question

true variation between centers

Solution

analyze by a random distribution of the center effects

ISCB00-page 13

Random center effects model

Assume

$$\theta_i \sim N(\mu, \tau^2)$$

Consequently

$$\hat{\theta}_i \sim N(\mu, s_i^2 + \tau^2)$$

Fit by ML

RESULTS

First step:

Test for homogeneity or is there is center effect in the data?

Homogeneity

either	$H_0 : \theta_1 = \theta_2 = \ldots\ldots \theta_{44} = \mu$
versus	$H_1 : \theta_1 \neq \theta_2 \neq \ldots\ldots \theta_{44} \neq \mu$
or	$H_0 : \tau^2 = 0$
versus	$H_1 : \tau^2 > 0$

fixed effect model	-2 LL= 126.7591	$\tau^2 = 0.2531$
random effect model	-2 LL= 118.2720	

LRT = 8.4 $(\chi^2_{(1/2)})$

Better estimate of the center effect

by Empirical Bayes

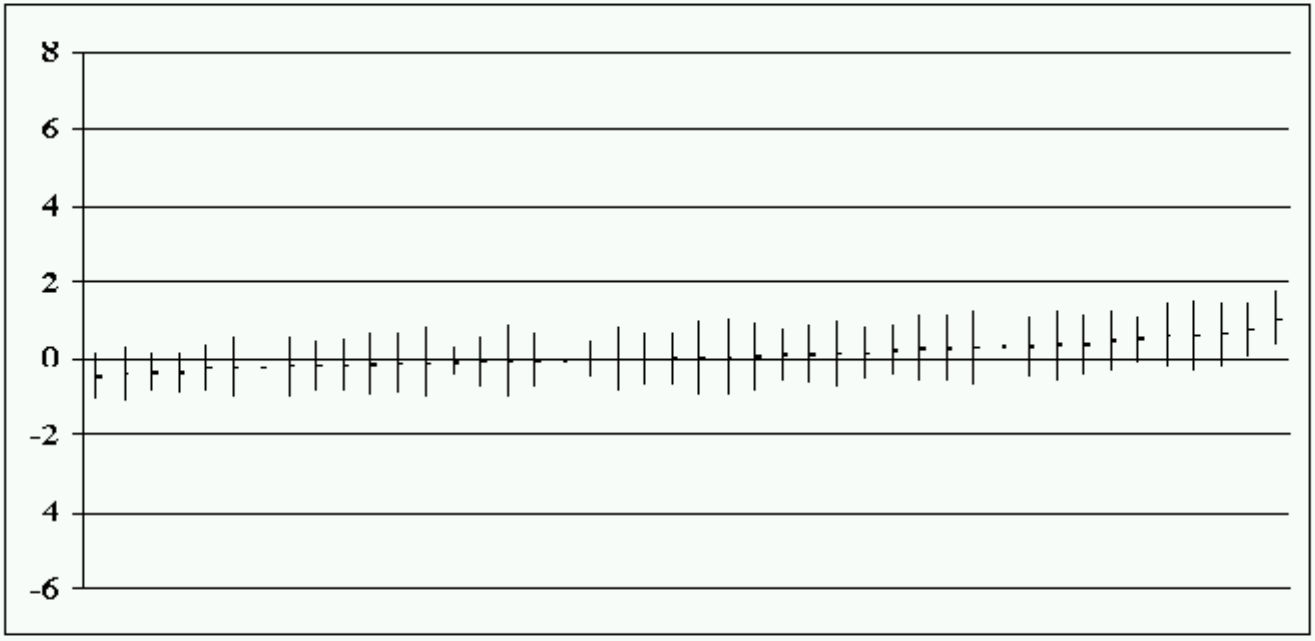
The posterior distribution of the center effect is given by

$$\theta_i \left| \hat{\theta}_i \sim N(EBE_i, pv_i) \right.$$

$$EBE_i = \mu + [\tau^2 / (\tau^2 + s_i^2)] * (\hat{\theta}_i - \mu)$$

$$pv_i = \tau^2 s_i^2 / (\tau^2 + s_i^2)$$

EBE center effects



$$EBE_i \pm 1.96\sqrt{pv_i}$$

ISCB00-page 17

Second step :

regression model for true θ_i's or
can center variability be explained ?

$$\theta_i = \beta_0 + \beta_1 Z_1 + \ldots + error$$
$$error \sim N(0, \tau^2)$$

	random effect variance	
	0.2531	
Factor	residual variance	explained variance
country	0.2191	19%
volume (cont's)	0.2321	8%
volume (dichtome)	0.1680	34%
volume and country	0.1438	43%

ISCB00-page 18

League tables or 'ranking' of centers

$$EPC_i = \phi[(EBE_i \quad \mu) / \sqrt{\tau^2 + pv_i}]$$

Measures of rankability/ true variation between centers

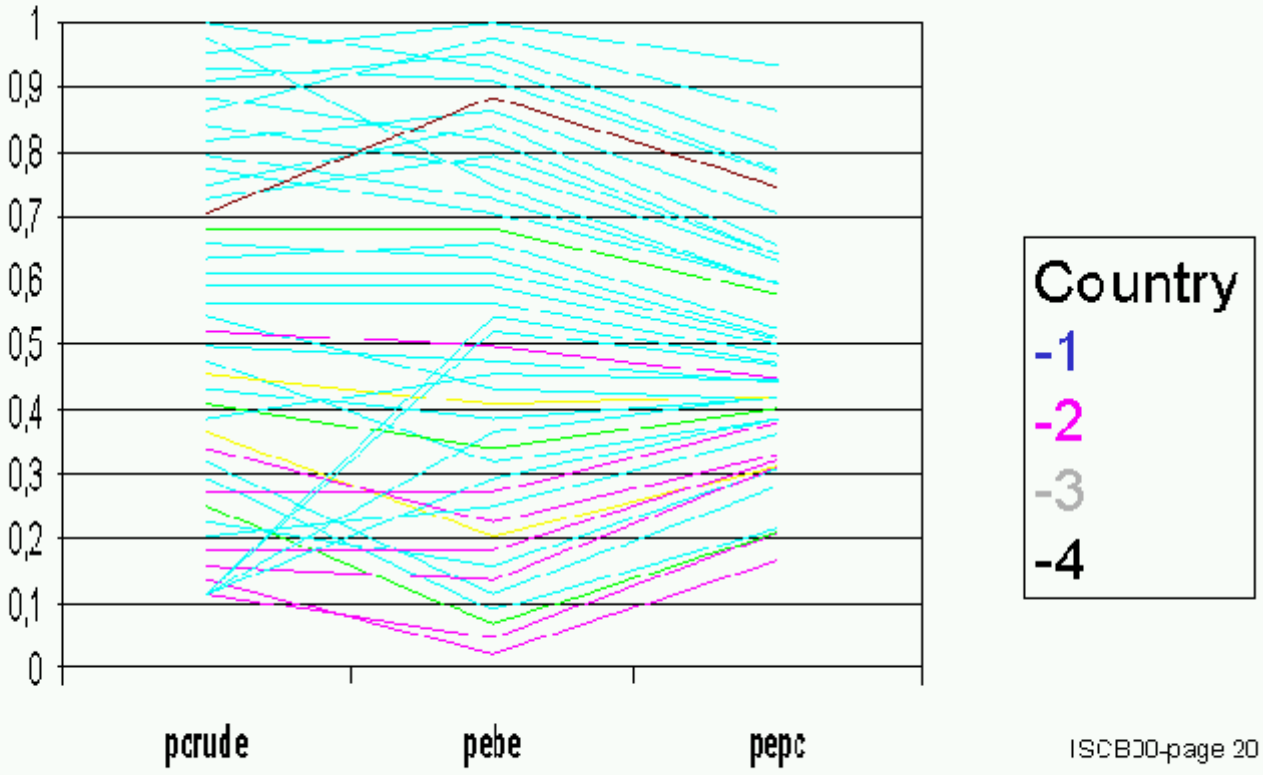
$$RA - 12 * \mathrm{var}(EPC)$$

$$R^2 = \tau^2 / [\tau^2 + median(s_i^2)]$$

Our data

RA = 40 %
R^2 = 41.6 %

Ranking based on raw data, EBE and EPC



ISCB30-page 20

conclusion

- score statistic approach for quality comparisons
- sensibility of league tables
- detection of outlying centers
- reference: Hans C van Houwelingen, Ronald Brand, Thomas A Louis. Empirical Bayes methods for monitoring health care quality. Bulletin of the International Statistical Institute. ISI '99 Conference Proceedings. Book 1, topic 7: 75 - 78.

Random effect model useful to study **goodness-of fit**

(le Cessie & van Houwelingen, 1995, Verweij and van Houwelingen, 1998)

General idea

- hazard for patient i $\ln(h_i(t)|X_i) = \ln(h_0(t)) + X_i\beta + u_i$
- $u'(u_1,...,u_n) \sim N(0, t^2 R)$
- test $t^2 = 0$ versus $t^2 > 0$
- test statistic $(d\&H)' R (d\&H) \& trace(RH)$

  under $t^2 = 0$ expectation=0, variance=.....

  corrections needed for estimation of ß and h_0

- Testing for cluster/centre effects by block-diagonal R.
- Suggestion of leCessie and Verweij, spatial R
- Next step: fitting “spline” to residuals

Demonstrated on Ovarian cancer data set.

No indication of lack of fit of the simple additive model

Quick and dirty method

approximate model without fitting linear model

- compute martingale residuals using Breslow hazard
- estimate effect per DIAM×FIGO×KARN combination
- fit model $\hat{?}_i \,'\, X_i ß \% u_i \% e_i$

  - e_i measurement error in $\hat{?}_i$
  - u_i deviation from model

Result

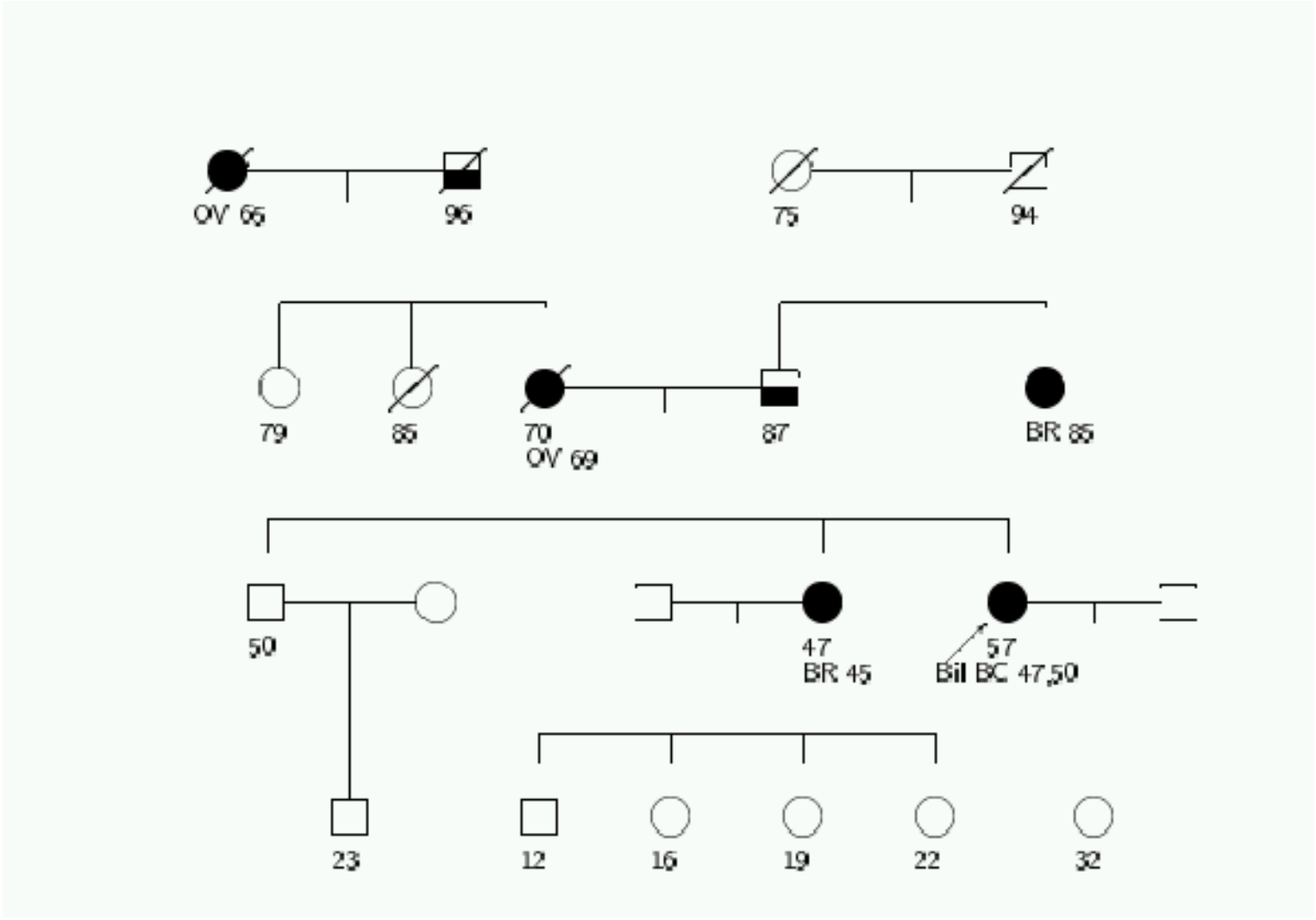
$\hat{var}(u) \,'\, 0$ regression-coefficients very close to formal Cox

Score functions in pedigree models

(Houwing-Duistermaat& Van Houwelingen, 1995, 2000)

- Pedigree-models
- Testing familiarity
- Predicting outcomes using family history

Pedigrees


OV 66
96
75
94
79
85
70
OV 69
87
BR 85
50
47
BR 45
57
Bil BC 47,50
23
12
16
19
22
32

general outcome $Y_i = X_i\beta + u_i + e_i$

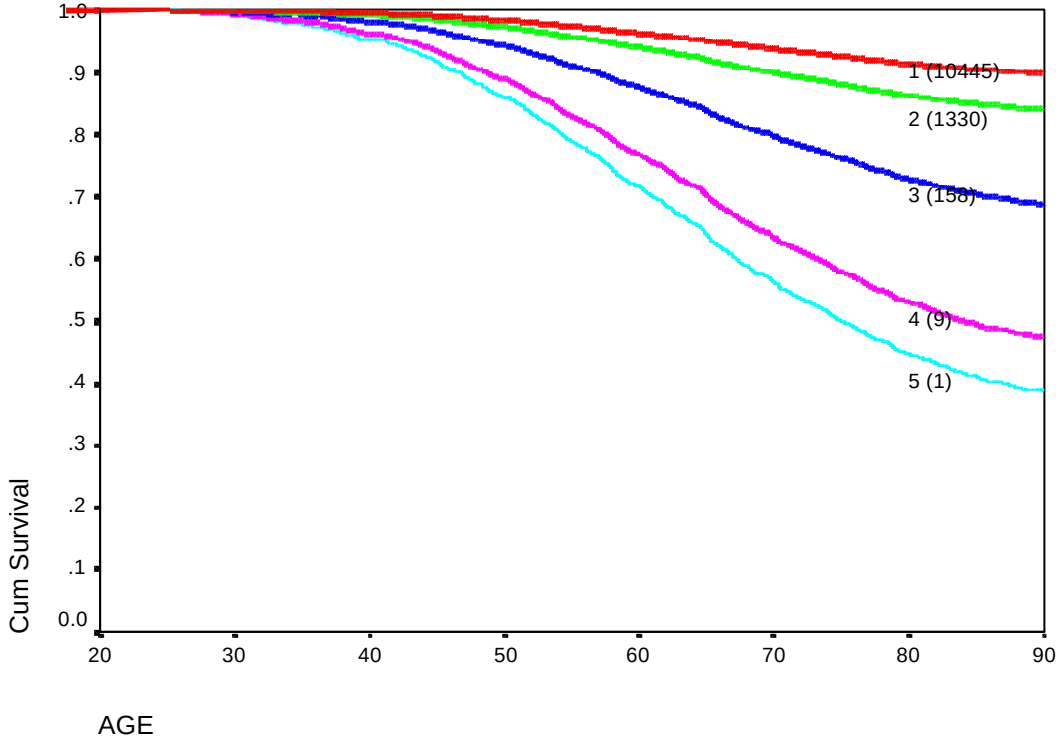
- u_i genetic part with genetic correlation $cov(u) = \sigma^2 R$
- $var(e_i) = \sigma^2$

Testing for familiarity = testing $t^2 = 0$ versus $t^2 > 0$

- General structure of test statistic $Q = \dfrac{(Y - X\hat{\beta})' R (Y - X\hat{\beta})}{\hat{\sigma}^2}$
- For survival structure as Verweij's goodness-of-fit test
- analysis of family history is regression on $R(d, H)$

In simulation study on breast cancer

- weighted sum of martingale residuals explains 80% of the explainable variation. but than number of affected relatives



stratified on # cases in the family

(..) gives the number in each stratum

General

- References in hand out

- Do not think survival analysis is very special.

  It is very close to just

  Generalised Linear (Mixed) Modeling