Analysis of Twin Data: Time to Event Models

J. Hjelmborg, T. Scheike, K. Holst and S. Møller

Epidemiology, Biostatistics and Biodemography, Institute of Public Health, University of Southern Denmark, Denmark

May 2015

Effect?

Exposure→Outcome

- Outcome: Time to occuence of event. Event may not occur - can be censored at follow-up.
- What is the contribution of genetic and environmental factors to the variation in risk of outcome?

$$\begin{cases} Y = \text{Genes} + \text{Environment} \\ \Sigma_Y = \Sigma_{\text{Genes}} + \Sigma_{\text{Environment}} \end{cases}$$

- What kind of genetic and environmental influences to expect?
- How does this influence vary with time?

Outline

- Introduction
- case-study: cancer diagnosis
- cumulative incidence.
- liability-threshold model extended.
- concordance and competing risks.
- practicals using R package 'mets'.

Time in Twin studies

- Suppose we're studying a dichotomous trait; Disease is present or not.
- Suppose data is complete in the sense that status of disease does not change anymore.
- Analysis: prevalence, concordance, correlation and biometric measures - Yes, We Can!
- Example: Stuttering in childhood (questionnaire answered by adults).
- -at least we do not hesitate to assume complete status.

Table: Genetic influence on Stuttering

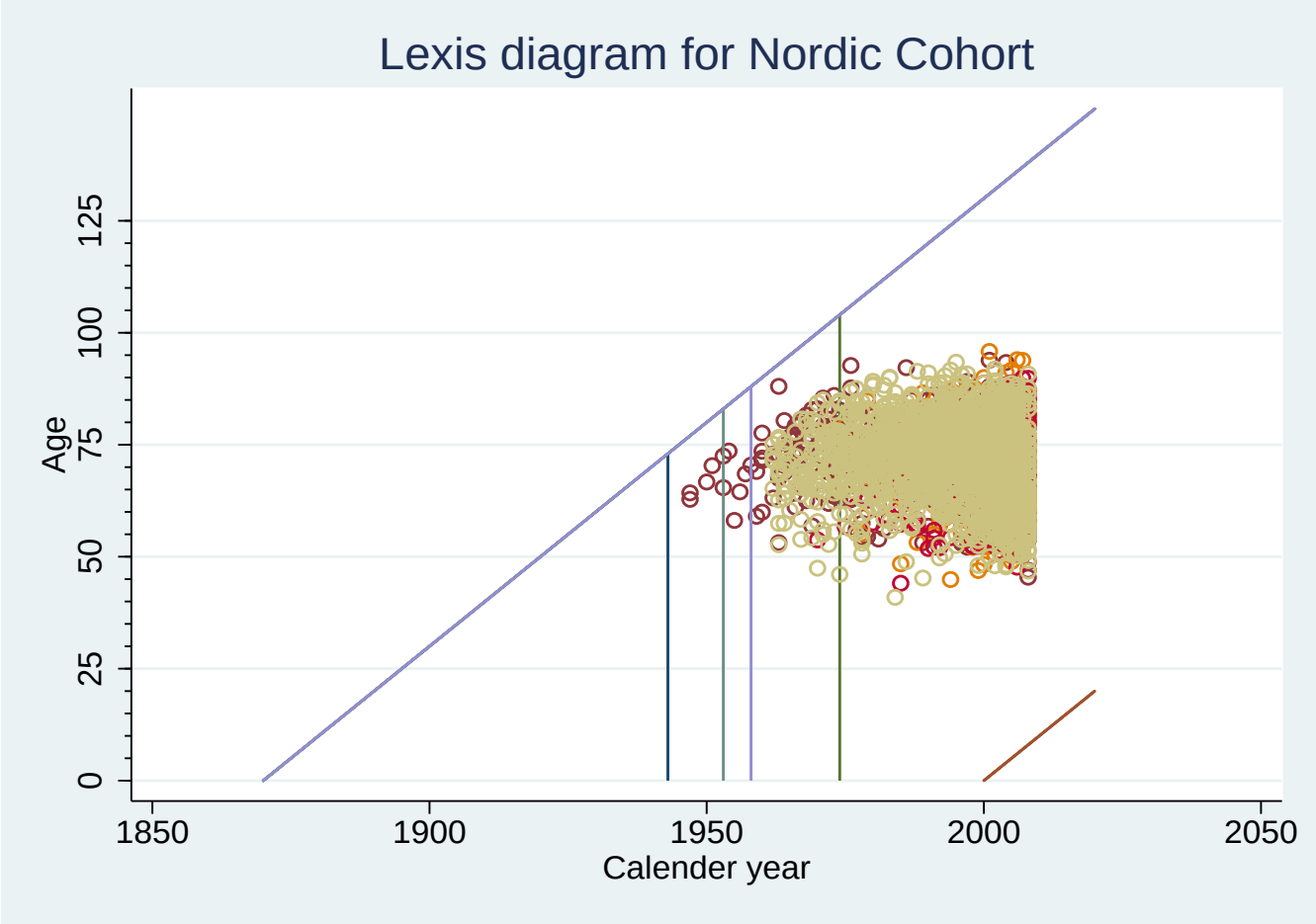
	Liability threshold model			
	prevalence	concordance	tetrachorics	heritability (95% CI)
MZ females	.04	.47 (.38,.59)	.81 (.71,.87)	.78 (.68,.85)
DZ females	.04	.08 (.04,.16)	.17 (-.02,.35)	AE model
MZ males	.08	.54 (.46,.62)	.79 (.72,.85)	.75 (.66,.82)
DZ males	.08	.10 (.062,.16)	.07 (-.07,.23)	AE model

(Fibiger et al. 2008)

Time in twin studies

- Fabulous Nordic data!
- Time: traits may change - hence results depend on time of observation.
- Can you think of a study, ie. trait and design, that is not governed by this?
- Fabulous Nordic data often contain registration of time of events!

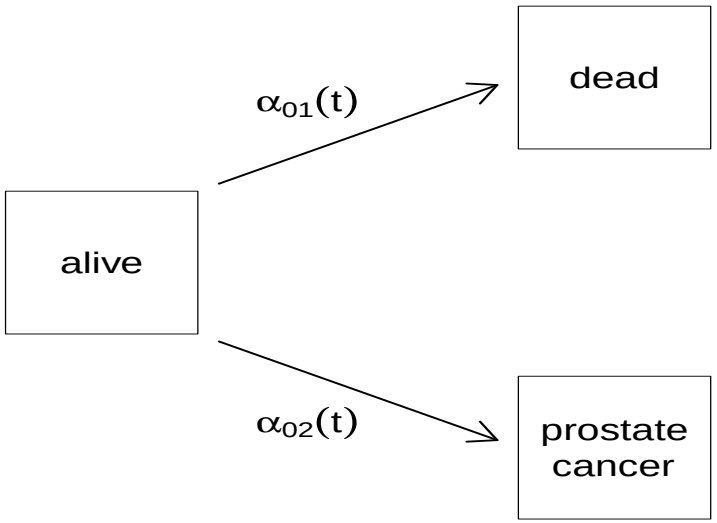
Lexis diagram - Nordic data on prostate cancer



- More than 70% are alive without cancer at follow-up.
- -also, delayed entry due to initiation of cancer registration.

Time in Twin studies

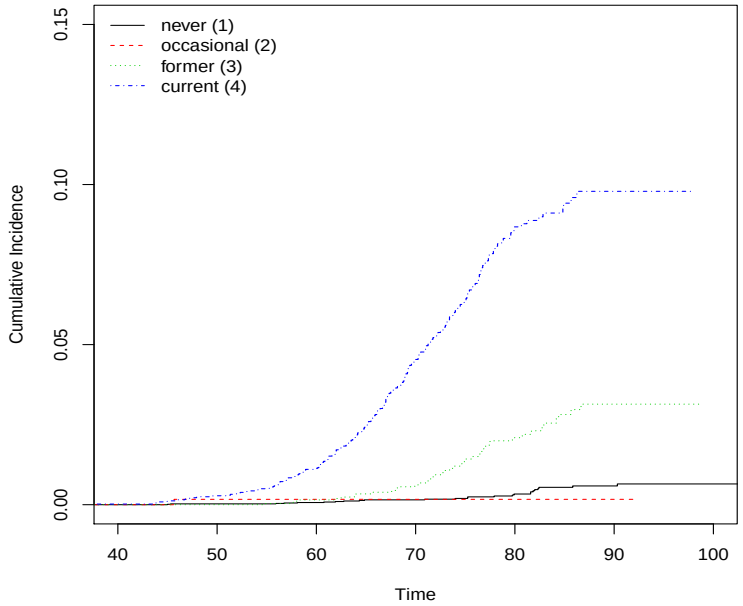
- We borrow methods from *survival analysis.*
- The Zoo: *events*, *censorings*, *competing risks*,. . .
- -a classic dichotomous trait is now an event.
- There may be multiple outcomes at each time point:

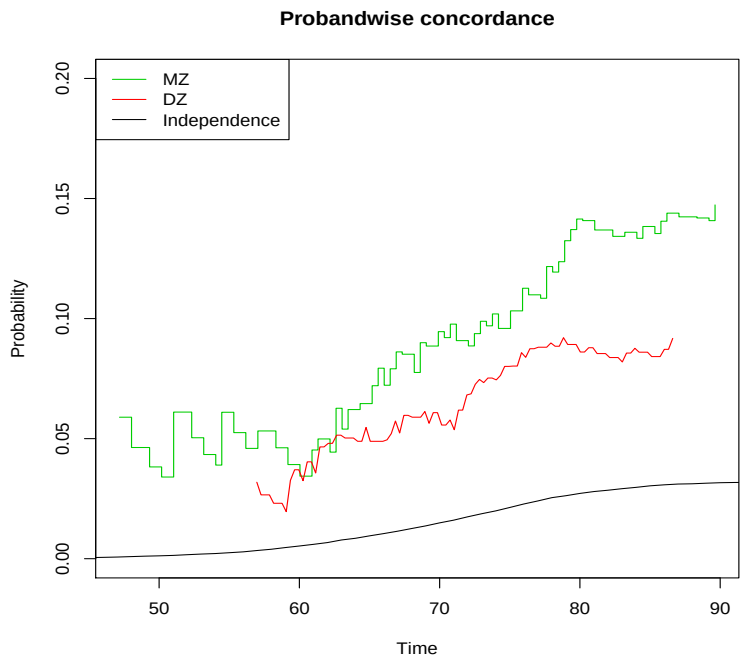


Time in Twin studies

Goals

- The cumulative incidence: *Risk of event before time t*
- The casewise concordance: *Risk of event in twin before time t given event in co-twin before time t*





Summary of sources of bias			
	prevalence	concordance	probandwise
All complete data (1)	biased (low or high)	biased	biased
All data (2)	too low	too low	biased
-and modelling censorings (3)	ok	ok	ok

1. In case (1) all complete data at follow-up is used, that is, censored data is excluded.
2. In case (2) all observed data is used including censored observations at follow up, that is, censored observations are ignored.
3. In case (3) censorings and competing events (eg. death before cancer) are modelled.

Sources of bias - breast cancer

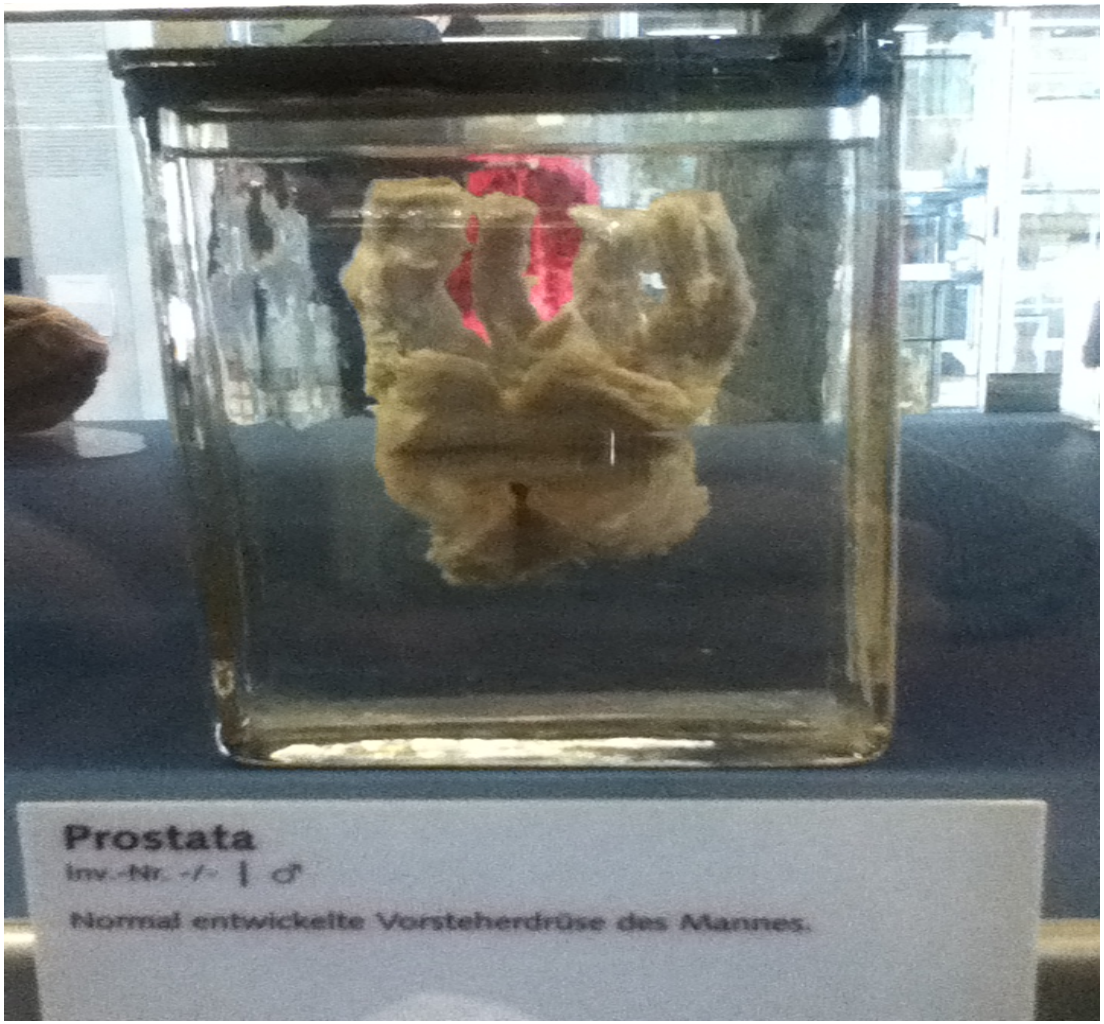
Summary of sources of bias

	prevalence	concordance	probandwise
All complete data (1)	biased (low or high)	biased	biased
All data (2)	too low	too low	biased
-and modelling censorings (3)	ok	ok	ok

Breast cancer risk and sources of bias

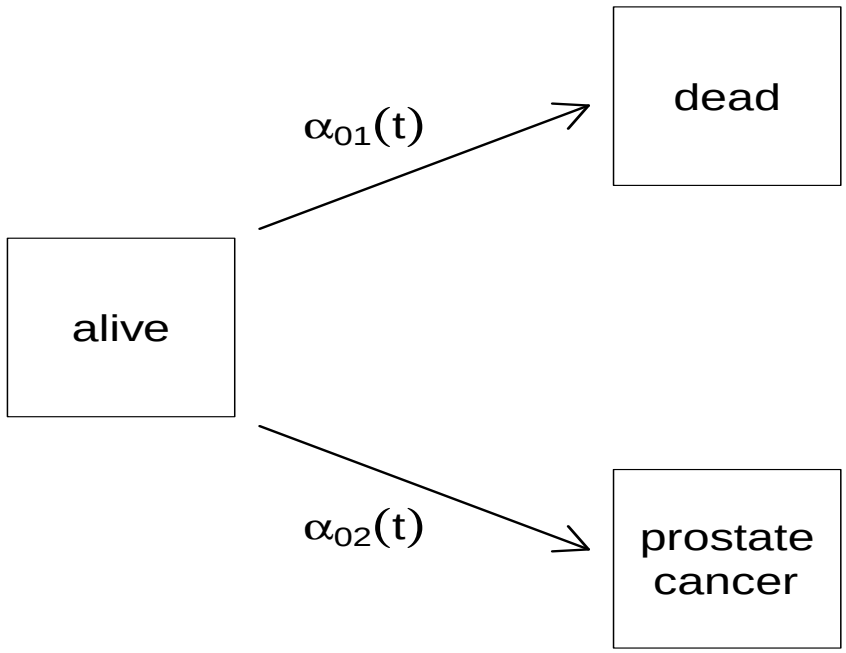
	Prevalence		Probandwise concordance	
	MZ twins	DZ twins	MZ twins	DZ twins
Complete data (1)	0.090 (0.005)	0.080 (0.004)	0.33 (0.04)	0.21 (0.03)
All data (2)	0.032 (0.002)	0.035 (0.001)	0.21 (0.03)	0.13 (0.02)
-and modelling censorings (3)	0.11 (0.004)	0.11 (0.004)	0.25 (0.04)	0.16 (0.03)

Example - Prostate cancer in twins



Methods - Competing risks

- 'the individual can experience more than one type of event?.
- 'when time to event is not independent of censoring-mechanism?.
- 'when other events precludes or interacts with event of interest?.



R Kioski - Package 'mets'

Prostate cancer in twins - Description


```
#Date: 2012-11-24

#Author: Klaus K. Holst, Thomas Scheike and
#        Jacob Hjelmborg
#Modified 2015-05-24

library(etm)

## Loading required package:  survival

library(prodlim)
library(mets)

## Loading required package:  timereg
## Warning:  package 'timereg' was built under R
version 3.1.3
## Loading required package:  lava
```

Prostate cancer in twins - Description


```
data(prt)  # simulated prostate cancer data
head(prt)

##     country     time status zyg id cancer
## 31 Denmark 96.98833      1  DZ  1      0
## 32 Denmark 80.88885      1  DZ  1      0
## 39 Denmark 68.04498      1  DZ  3      0
## 40 Denmark 61.45903      1  DZ  3      0
## 51 Denmark 78.78068      1  DZ  5      0
## 52 Denmark 90.36252      1  DZ  5      0
```

Prostate cancer in twins - Description


```
kable(with(prt, table(status,country)))
```

	Denmark	Finland	Norway	Sweden
0	7300	2533	3102	8348
1	2223	1209	876	2689
2	148	184	129	481

Prostate cancer in twins - Description


```
kable(with(prt, table(cancer,zyg)))
```

	DZ	MZ
0	17408	10872
1	583	359

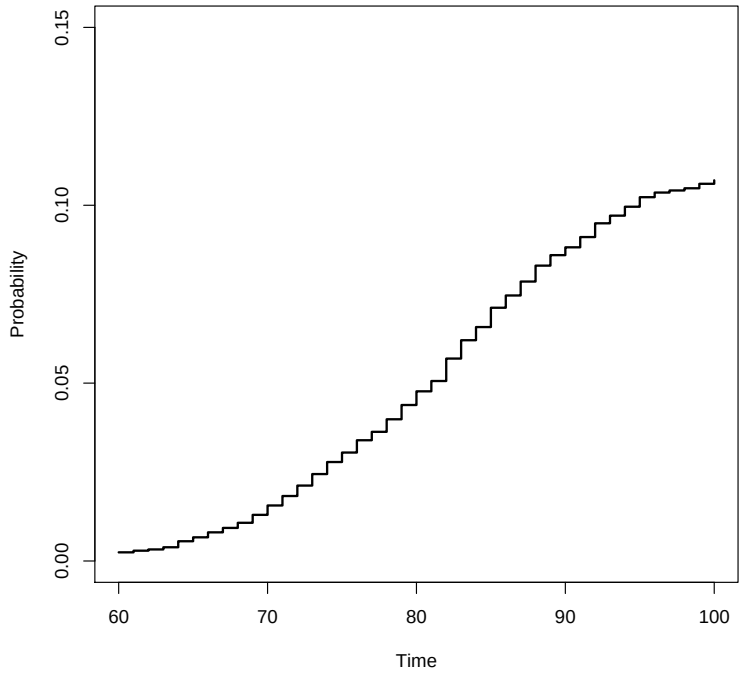
```
out <- lm(cancer~-1+zyg,prt)  # lifetime risk (!).
kable(summary(out)$coef, digits=2)
```

	Estimate	Std. Error	t value	Pr(>|t|)
zygDZ	0.03	0	24.61	0
zygMZ	0.03	0	19.18	0

Prostate cancer in twins - cumulative incidence


```
times <- seq(60,100,by=1)  # set time-range - elderly males
cifmod <- comp.risk(Event(time,status)~+1+cluster(id),data=prt,cause=2,n.sim=0,
                    times=times,conservative=1,max.clust=NULL,model="fg")
pcif <- predict(cifmod,X=1,resample.iid=0,uniform=0,se=0)
```

```
plot(pcif,multiple=1,se=0,uniform=0,ylim=c(0,0.15))
```



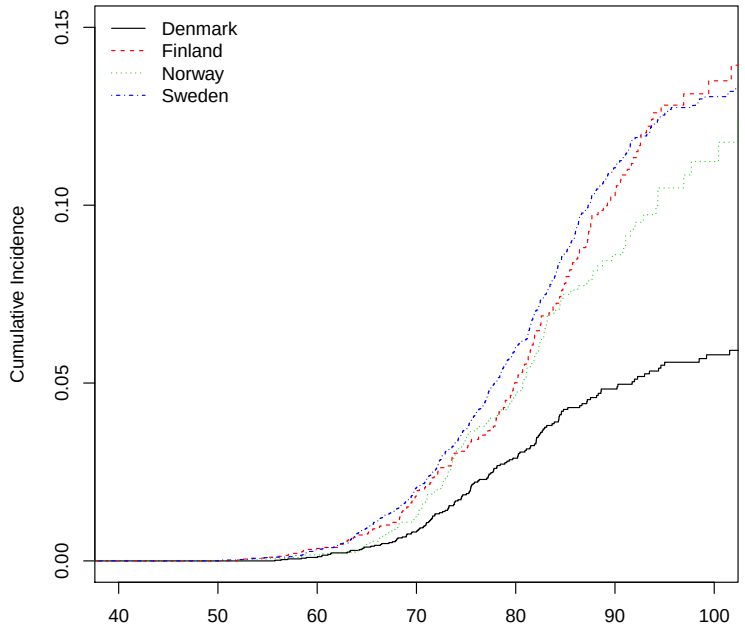
Prostate cancer in twins - cross odds ratio


```
theta.des <- model.matrix(~-1+factor(zyg),data=prt) ## design for MZ/DZ status
or1 <- or.cif(cifmod,data=prt,cause1=2,cause2=2,theta.des=theta.des,
              same.cens=TRUE, score.method="fisher.scoring")  # see help(or.cif)
summary(or1)


## OR for dependence for competing risks
##
## OR of cumulative incidence for cause1= 2  and cause2= 2
##                log-ratio Coef.    SE    z    P-val Ratio    SE
## factor(zyg)DZ            0.80 0.221 3.61 3.01e-04  2.22 0.492
## factor(zyg)MZ            2.09 0.276 7.56 4.13e-14  8.07 2.230


or1$score


##              [,1]
## [1,] 6.417230e-09
## [2,] 1.197999e-07
```

Prostate cancer in twins - cumulative incidence


```
# By non-parametric Aalen-Johansen estimator.
addprtmenetmd<-etmCIF(Surv(time,status!=0)~+factor(country),
                      data=prt,etype=status,failcode=2)
```

```
plot(addprtmenetmd,ylim=c(0,0.15),col=1:4,xlim=c(40,100),
     curvlab = c("Denmark", "Finland", "Norway","Sweden"))
```



Prostate cancer in twins - concordance


```
### ignoring country
### marginal cumulative incidence of prostate cancer##'
outm <- prodlim(Hist(time,status)~+1,data=prt)

times <- 60:100
cifmz <- predict(outm,cause=2,time=times,newdata=data.frame(zyg="MZ")) ## cause is 2 (second cau
cifdz <- predict(outm,cause=2,time=times,newdata=data.frame(zyg="DZ"))

### concordance for MZ and DZ twins
cc <- bicomprisk(Event(time,status)~strata(zyg)+id(id),data=prt,cause=c(2,2),prodlim=TRUE)


## Strata 'DZ'
## Strata 'MZ'

cdz <- cc$model$"DZ"
cmz <- cc$model$"MZ"

cdz <- casewise(cdz,outm,cause.marg=2)
cmz <- casewise(cmz,outm,cause.marg=2)
```

Prostate cancer in twins - concordance


```
plot(cmz,ci=NULL,ylim=c(0,0.6),xlim=c(60,100),legend=TRUE,col=c(3
par(new=TRUE)
plot(cdz,ci=NULL,ylim=c(0,0.6),xlim=c(60,100),legend=TRUE)
```

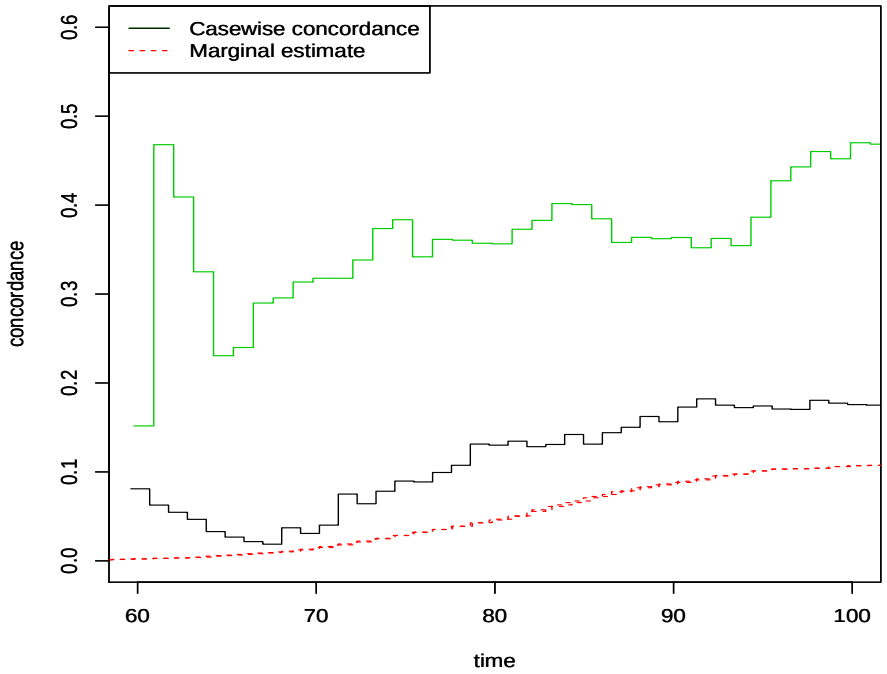


Figure: Casewise concordance

Prostate cancer in twins - Concordance

More information from

```
summary(cmz)
summary(cdz)
```

Further, Relative recurrence risk, multiple locus index and other measures can be obtained.

Time to event - biometric modeling?

The New England
Journal of Medicine

VOLUME 343 JULY 13, 2000 NUMBER 2

ENVIRONMENTAL AND HERITABLE FACTORS IN THE CAUSATION OF CANCER

Analyses of Cohorts of Twins from Sweden, Denmark, and Finland

PAUL LICHTENSTEIN, PH.D., NIELS V. HOLM, M.D., PH.D., PIA K. VERKASALO, M.D., PH.D., ANASTASIA ILIADOU, M.SC., JAAKKO KAPRIO, M.D., PH.D., MARKKU KOSKENVUO, M.D., PH.D., EERO PUKKALA, PH.D., AXEL SKYTTHE, M.SC., AND KARI HEMMINKI, M.D., PH.D.

- NEJM 2000 landmark paper report heritabilities for all cancer sites.
- Prostate cancer: case-wise concordance rates (MZ; DZ) of 0.20; 0.09, and a heritability of 0.42 (0.29; 0.50).
- Biometric model: Liability threshold (ignoring censored data, $\sim$70%).
- Let's take censoring into account - Aim for NorTwinCan Study.

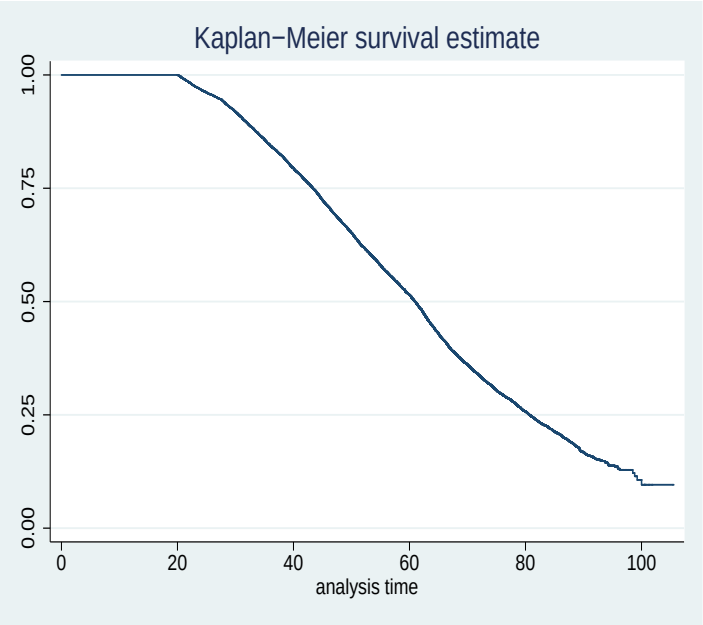
Time to event - biometric modeling - a first attempt

Genetic influence on risk scale, how about heritability?

- Liability-threshold polygenic ADCE model.:

$$\text{probit}(P(\text{twin j gets cancer}|X_j, Z)) = X_j^T \beta + Z, \quad j = 1, 2$$

- Extension: Weights from inverse probability of censoring:



Liability threshold model with IPW

- Liability model with Inverse Probability Weighting and adjusting for covariates
- Probabilities of being censored - we weight complete observations with these. In analogy with missing data analysis assuming missing at random (MAR). Probability weights based on Aalen's additive model

Liability threshold: Eq. marginals for twins


```
bp.flex <- twinlm.time(cancer~country,zyg="zyg",
    DZ="DZ",id="id",
    cumulative = TRUE, binary=TRUE,
    type="flex",data=prt,
    cens.formula=Surv(time,status==0)~1+zyg+country,
    breaks=Inf,
       control=list(refit=TRUE))
round(summary(bp.flex)$coef,2)

##                              Estimate Std.Err 2.5% 97.5%
## Tetrachoric correlation MZ       0.70    0.05 0.58  0.78
## Tetrachoric correlation DZ       0.27    0.06 0.14  0.39
```

Liability threshold: Eq. marginals for twins MZ and DZ


```
bp.u <- twinlm.time(cancer~country,zyg="zyg",
    DZ="DZ",id="id",
    cumulative = TRUE, binary=TRUE,
    type="u",data=prt,
    cens.formula=Surv(time,status==0)~1+zyg+country,
    breaks=Inf,
       control=list(refit=TRUE))

round(summary(bp.u)$coef,2)
```

```
##                              Estimate Std.Err 2.5% 97.5%
## Tetrachoric correlation MZ       0.69    0.05 0.58  0.78
## Tetrachoric correlation DZ       0.28    0.07 0.14  0.40
```

Liability threshold: Eq. marginals for twins MZ and DZ

We can compare above models directly since nested:

```
compare(bp.u,bp.flex)

##
##  - Likelihood ratio test -
##
## data:
## chisq = 19.8361, df = 4, p-value = 0.000538
## sample estimates:
## log likelihood (model 1) log likelihood (model 2)
##                -8323.003                -8313.085
```

Liability threshold: ACE with IPW


```
bp.ace <- twinlm.time(cancer~country,zyg="zyg",
    DZ="DZ",id="id",
    cumulative = TRUE, binary=TRUE,
    type="ace",data=prt,
    cens.formula=Surv(time,status==0)~1+zyg+country,
    breaks=Inf,
              control=list(refit=TRUE))

score(bp.ace)

## [1]  1.089056e-04  3.706803e-05  2.993500e-05 -2.206747e-06  8

round(summary(bp.ace)$coef,2)

##                     Estimate Std.Err 2.5% 97.5%
## A                       0.67    0.05 0.58  0.77
## C                       0.00    0.00 0.00  0.00
## E                       0.33    0.05 0.23  0.42
## MZ Tetrachoric Cor      0.67    0.05 0.56  0.76
```

Liability threshold: ADE with IPW


```
bp.ade <- twinlm.time(cancer~country,zyg="zyg",
    DZ="DZ",id="id",
    cumulative = TRUE, binary=TRUE,
    type="ade",data=prt,
    cens.formula=Surv(time,status==0)~1+zyg+country,
    breaks=Inf,
       control=list(refit=TRUE))

round(summary(bp.ade)$coef,2)

##                     Estimate Std.Err  2.5% 97.5%
## A                       0.42    0.27 -0.11  0.95
## D                       0.27    0.28 -0.29  0.83
## E                       0.31    0.05  0.21  0.41
## MZ Tetrachoric Cor      0.69    0.05  0.58  0.78
## DZ Tetrachoric Cor      0.28    0.07  0.14  0.40
```

Liability threshold: ACE versus ADE

We can compare above models via the Akaike Information Index:

```
AIC(bp.ace, bp.ade)

##        df      AIC
## bp.ace  6 16662.82
## bp.ade  6 16658.01
```

Liability threshold: Stratified analysis


```
bp.ace.strata <- twinlm.time(cancer~strata(country),zyg="zyg",
    DZ="DZ",id="id",
    cumulative = TRUE, binary=TRUE,
    type="ace",data=prt,
    cens.formula=Surv(time,status==0)~1+zyg+country,
    breaks=Inf,
       control=list(refit=TRUE))

## Strata 'Denmark'
## Strata 'Finland'
## Strata 'Norway'
## Strata 'Sweden'

summary(bp.ace.strata)

## ------------------------------
## Strata 'Denmark'
## ------------------------------
## Strata 'Finland'
## ------------------------------
```

Liability threshold: Cumulative heritability


```
bp.ace.cum <- twinlm.time(cancer~country,zyg="zyg",
    DZ="DZ",id="id",
    cumulative = TRUE, binary=TRUE,
    type="ace",data=prt,
    cens.formula=Surv(time,status==0)~1+zyg+country,
    breaks=seq(60,90, by=2),
       control=list(refit=TRUE))
names(bp.ace.cum)
bp.ace.cum$summary
summary(bp.ace.cum)
```

Prostate cancer in twins - casewise concordance


```
plot(bp.ace.cum,which=c(8,13),ylim=c(0,0.5),legendpos="topright",
     col=c("darkred","darkblue"),lty=c(1,2),
     legend=c("MZ and 95% CI.","DZ and 95% CI."),
     ylab="Casewise concordance") #,    main="Nordic twin cohorts
```

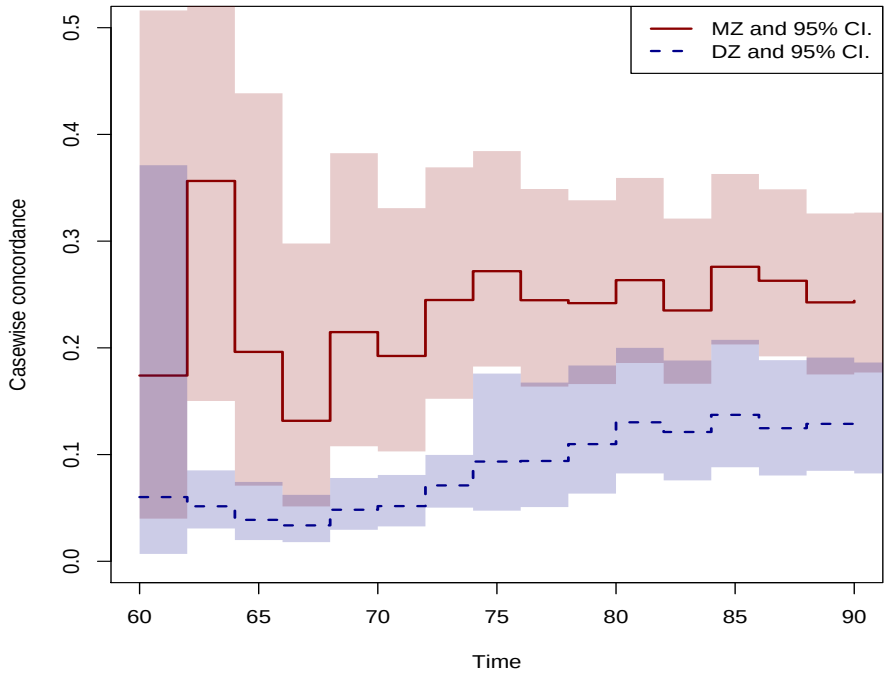


Figure: Casewise concordance

Prostate cancer in twins - heritability


```
plot(bp.ace.cum,which=c(1),ylim=c(0,1),legendpos="bottomright",
     col=c("darkred"),lty=c(1),
     legend=c("H2 and 95% CI."),
     ylab="Heritability") #,    main="Nordic twin cohorts")
```

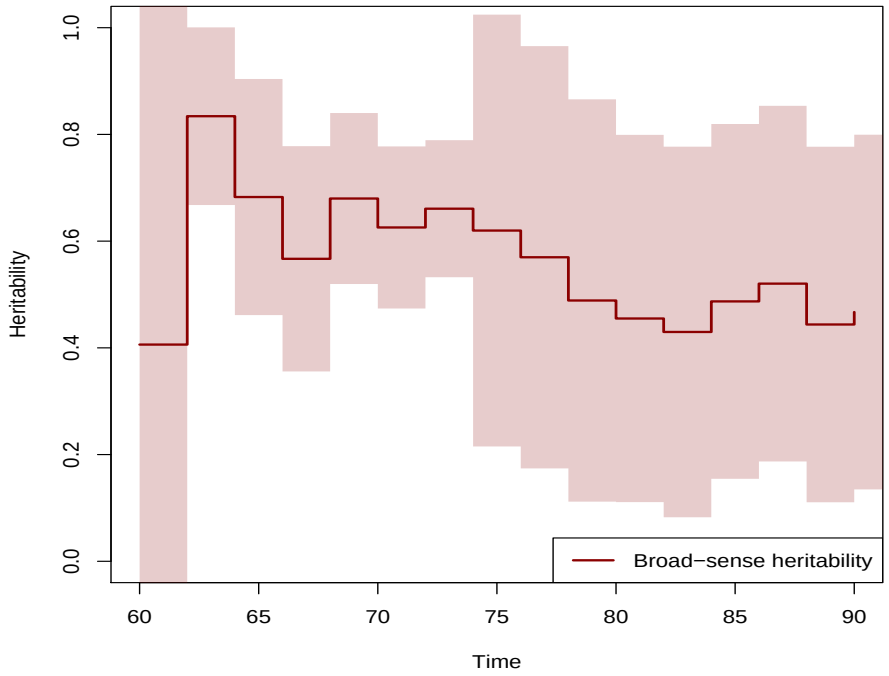


Figure: Cumulative heritability

Exercise

- Create above plots of cumulative casewise concordance and heritability from the liability threshold ADE model with IPW for censoring.
- What does the above stratified analysis add?

Exercise

- What would happen if time to event was ignored?
- This can be investigated by repeating the analysis without IPW.
- See the following slides for implementation.

Liability threshold: Saturated model - ignoring time


```
bp0 <- biprobit(cancer~country + cluster(id)+strata(zyg),
                data=prt)


## Strata 'DZ'
## Strata 'MZ'

summary(bp0)

## ------------------------------
## Strata 'DZ'
## ------------------------------
## Strata 'MZ'
```

Liability threshold: Eq. marginals - ignoring time


```
bp1 <- bptwin(cancer~country,zyg="zyg",DZ="DZ",id="id", binary=TRUE, type="u",data=prt)
summary(bp1)$probMZ
```

```
##                          Estimate         2.5%        97.5%
## Concordance           0.004467324  0.003292577  0.006058658
## Casewise Concordance  0.293233128  0.234208848  0.360136998
## Marginal              0.015234718  0.012860807  0.018038809
## Rel.Recur.Risk       19.247690103 14.645257725 23.850122482
## log(OR)               3.625153734  3.234286444  4.016021025
```

```
summary(bp1)$probDZ
```

```
##                          Estimate         2.5%       97.5%
## Concordance           0.001440254 0.0009503536 0.002182143
## Casewise Concordance  0.094537629 0.0667986510 0.132164199
## Marginal              0.015234718 0.0128608072 0.018038809
## Rel.Recur.Risk        6.205407284 4.0723812095 8.338433359
## log(OR)               1.994581307 1.5853517140 2.403810901
```

```
summary(bp1)$coef
```

```
##                              Estimate    Std.Err      2.5%     97.5%
## Tetrachoric correlation MZ 0.6988528 0.03375873 0.6265551 0.7592258
## Tetrachoric correlation DZ 0.3706259 0.04339034 0.2826528 0.4524161
```

```
compare(bp0,bp1) # LRT
```

Liability threshold: ACE model - ignoring time


```
bp2 <- bptwin(cancer~country,zyg="zyg",DZ="DZ",id="id", binary=TRUE,type="ace",data=prt)
summary(bp2)$probMZ
```

```
##                          Estimate         2.5%        97.5%
## Concordance           0.004467383  0.003292626  0.006058727
## Casewise Concordance  0.293234795  0.234210425  0.360138679
## Marginal              0.015234832  0.012860918  0.018038923
## Rel.Recur.Risk       19.247655571 14.645254708 23.850056434
## log(OR)               3.625156484  3.234289702  4.016023266
```

```
summary(bp2)$probDZ
```

```
##                         Estimate         2.5%       97.5%
## Concordance          0.00144021 0.0009503172 0.002182094
## Casewise Concordance 0.09453405 0.0667954859 0.132160423
## Marginal             0.01523483 0.0128609183 0.018038923
## Rel.Recur.Risk       6.20512575 4.0721465918 8.338104900
## log(OR)              1.99452774 1.5852906534 2.403764836
```

```
summary(bp2)$coef
```

```
##                        Estimate    Std.Err       2.5%     97.5%
## A                   0.65647764 0.10956971  0.4417250 0.8712303
## C                   0.04237639 0.09289080 -0.1396862 0.2244390
## E                   0.30114597 0.03375863  0.2349803 0.3673117
## MZ Tetrachoric Cor 0.69885403 0.03375863  0.6265565 0.7592268
## DZ Tetrachoric Cor 0.37061521 0.04339116  0.2826405 0.4524070
```

The Heritability of Prostate Cancer in the Nordic Twin Study of Cancer Hjelmborg, Scheike, Kaprio, Mucci et al.
`http://cebp.aacrjournals.org/content/23/11/2303`
Cancer Epidemiology, Biomarkers & Prevention (2014)

References

- *Estimating heritability for cause specific mortality based on twin studies*; Scheike, Holst and Hjelmborg; *LIDA* (2013).
- *Estimating twin concordance for bivariate competing risks twin data*; Scheike, Holst and Hjelmborg; *Stat Med* (2014)
- *Measuring early or late dependence for bivariate lifetimes of twins* Scheike, Holst and Hjelmborg; *LIDA* (2014).
- *Revisiting the Concordance for Twin Pairs*; Hjelmborg, Scheike, Holst and Möller; *Hum Genet Twin Research* (2015), in preparation
- *The liability threshold model for censored twin data* Holst, Scheike and Hjelmborg, *Computational Statistics & Data Analysis* (2015)

Cumulative incidence adjusted for censoring

- Data: $(X_{jk}, \tilde{T}_{jk}, \tilde{\epsilon}_{jk})$, where $\tilde{T}_{jk} = \min(T_{jk}, C_k)$ for $j \in \{1, 2\}$, $k \in \{1, \ldots, p\}$.
- The cumulative incidence function:

$$F_1(t) = \text{Prob}(T \leq t, \epsilon = 1) = \int_0^t \lambda_1(s) S(s-) ds,$$

- Non-parametric Aalen-Johansen product-limit mle estimator; $\hat{F}_1(t) = \sum_{t_j \leq t} \frac{d_{ij}}{n_j} \hat{S}(t_{j-1})$
- Models: Fine and Gray (1999) 'proportional model?, also logistic link and random effects extension.
- Methods: *Estimation of probandwise concordance for censored time to event twin data*. Scheike, Holst and Hjelmborg (2012).

- Methods: *Estimation of probandwise concordance for censored time to event twin data*. Scheike, Holst and Hjelmborg (2012).
- First, advanced model approach and then shortcut assuming same censoring in pairs.
- Warning: offensive slides may occur.

Concordance

- The concordance function conditional on covariates

$$C_X(t) = P_{j_1,j_2}(t,t) = P(T_1 \leq t, \epsilon_1 = j_1, T_2 \leq t, \epsilon_2 = j_2 | X),$$

- assume a specific semi-parametric regression structure

$$C_X(t) = h_c(\Lambda_0(t), X\beta_c, t)$$

- For example $h(a,b,t) = 1 - \exp(ab)$, or a logit-link version where $h(a,b,t) = (\exp(a+b)/(1+\exp(a+b))$. Here $\Lambda_0(t)$ is an increasing baseline-function.
- -in this way, semi-parametric random effects regression model (Scheike et al. 2012).

Methods - Competing risk random effects model

- Katsahian et al. (2006), Scheike et al. (2010):

$$F_1(t, X_j, Z) = P(T \leq t, \text{cancer} | Z, X_j) = 1 - \exp\left(-Z\Psi_\theta^{-1}\left[\Lambda(t, X_j)\right]\right)$$

  - $\Psi_\theta(t) = E_\theta\{\exp(-Zt) \mid X\}$ is the Laplace transform of the random effect Z.
  - Z is assumed gamma distributed with shape parameter $1/\theta$ and scale parameter θ.
  - $\Lambda(t, X_j) = \exp(-\alpha(t) - (\gamma X_j)t)$, where $\alpha(t)$ is a baseline regression function and γ is a q-dimensional vector of parameters.
  - This leads to random effects with mean 1 and variance θ.
- The marginal transition probability given only X_j is of the simple additive form

$$F_1(t, X_j) = 1 - \exp(-\alpha(t) - tX_j^T\beta)$$

Methods - Competing risk random effects model

- The marginal transition probability given only X_j is of the simple additive form

$$F_1(t, X_j) = 1 - \exp(-\alpha(t) - tX_j^T\beta)$$

- NB! If no covariates, $F_1(t)$ is simply a reparametrization of the non-parametric cumulative incidence function.
- For event times (T_1, T_2) we can compute the bivariate cumulative incidence $P(T_1 \leq t, \epsilon_1 = 1, T_2 \leq s, \epsilon_2 = 1|X_1, X_2)$ by

$$1-(1-F_1(t, X_1))-(1-F_1(s, X_2))+\Psi_\theta\left(\Psi_\theta^{-1}\left[\Lambda(t, X_1)\right] + \Psi_\theta^{-1}\left[\right.\right.$$

- Variances of random effects measures amount of correlation within pairs.

Concordance - much simpler!

- The concordance function conditional on covariates

$$C_X(t) = P_{j_1,j_2}(t,t) = P(T_1 \leq t, \epsilon_1 = j_1, T_2 \leq t, \epsilon_2 = j_2 | X)$$

- First, consider times to the joint event (j_1, j_2).
- When does the counting processes

$$N_{j_1,j_2}(t) = I(T_1 \leq t, \epsilon_1 = j_1, T_2 \leq t, \epsilon_2 = j_2) \quad \text{unobserved co}$$
$$\tilde{N}_{j_1,j_2}(t) = I(\tilde{T}_1 \leq t, \epsilon_1 = j_1, \tilde{T}_2 \leq t, \epsilon_2 = j_2) \quad \text{observed data}$$

have the same intensity function $\lambda_{j_1,j_2}(t)$?

Most important slide!

The intensity of the observed counting process equals that of the unobserved complete data process when

- censoring is the same in pairs and
- censoring is independent of competing risks data given covariates.

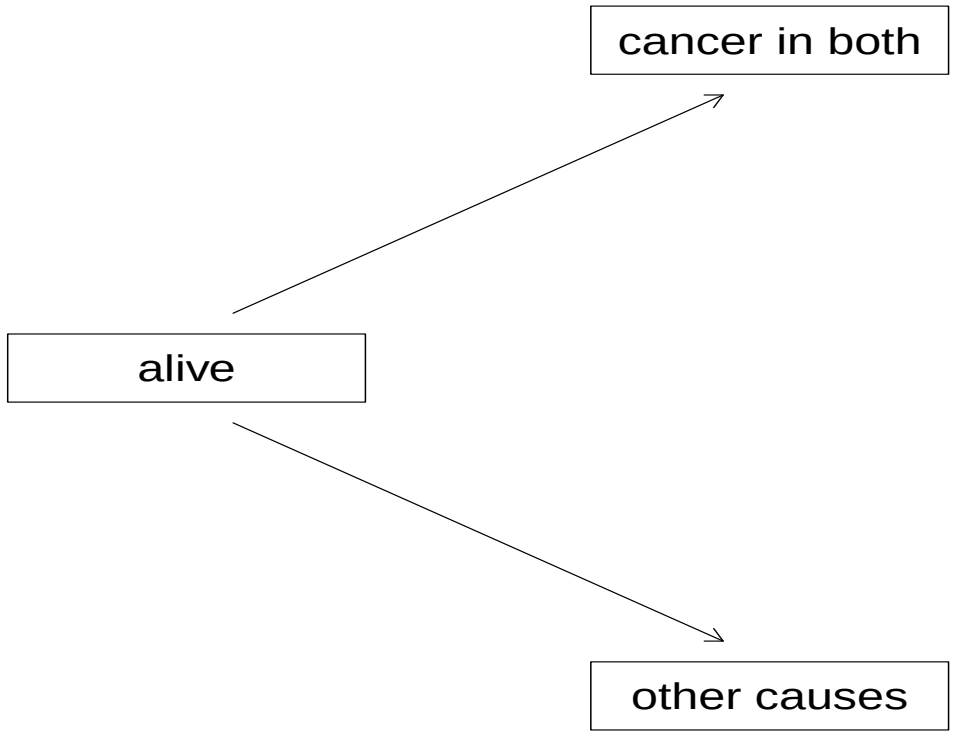
Formally,

$$P(\Delta\tilde{N}_{j_1,j_2}(t) = 1 | Y_{j_1,j_2}(t) = 1, X) =$$
$$= P(\Delta\tilde{N}_{j_1,j_2}(t) = 1 | \tilde{A}_1(t), \tilde{A}_2(t), (\tilde{T}_2 \geq t, \tilde{T}_1 \geq t), X)$$
$$= P(\Delta\tilde{N}_{j_1,j_2}(t) = 1 | \tilde{A}_1(t), \tilde{A}_2(t), (\tilde{T}_2 \geq t, \tilde{T}_1 \geq t), C > t,$$
$$= P(\Delta N_{j_1,j_2}(t) = 1 | A_1(t), A_2(t), (T_2 \geq t, T_1 \geq t), C > t,$$

with $\tilde{A}_1(t) = (\tilde{T}_1 \leq t, \epsilon_1 = j_1, \tilde{T}_2 \geq t)$ and
$\tilde{A}_2(t) = (\tilde{T}_2 \leq t, \epsilon_2 = j_2, \tilde{T}_1 \geq t)$.

Methods - applying same censoring

- Same censoring is very often plausible - observations at almost same time.
- We may then consider pairwise competing risk data.
- This is univariate - hence standard methods are applicable.



Reducing data to pairwise competing risks data that can be used for standard estimation.

twins	time	cause/status
twin 1 pair 1	10	cancer
twin 2 pair 1	20	cancer
combined data	20	cancer
twin 1 pair 2	10	cancer
twin 2 pair 2	20	dead
combined data	20	other causes
twin 1 pair 3	10	dead
twin 2 pair 3	20	cancer
combined data	10	other causes
twin 1 pair 4	10	dead
twin 2 pair 4	20	censored
combined data	10	other causes
twin 1 pair 5	10	cancer
twin 2 pair 5	20	censored
combined data	20	censored