Epidemiologic Data Analysis using R
Part 9
Time dependent covariates

Janne Pitkäniemi

Tampere University, Faculty of Social Sciences

University of Tampere Postgraduate training

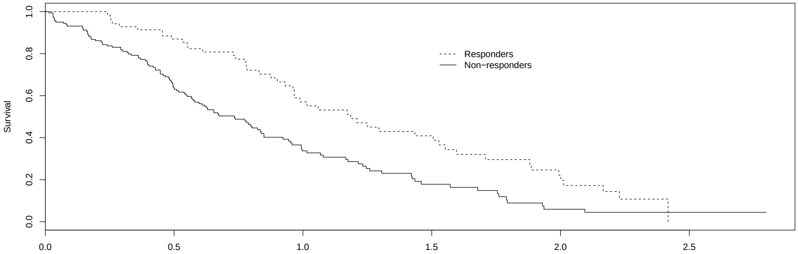
Feb 26 - Apr 9, 2018

Time-Dependent covariates in TTE

- A key underlying condition (a martingale) is that present actions depend only on the past.
- The decision of whether one is in the risk set or not and covariates can depend in any way on prior covariates and risk set patterns
- *but cannot look into the future.*

Analysis by treatment response - NO ! (T. Thernau)

- An example based on the advanced lung cancer data set, where assignment of the responder was made by looking
- The problem arises because any early deaths, those that occur before response can be assessed, will all be assigned to the non-responder group, even deaths that have nothing to do with the condition under study.
- Assume that subjects came in monthly for 12 cycles of treatment, and randomly declare a **response** for 5% of the subjects at each visit.



Classification of time dependent covariates

- Values change continuosly (change in small during a small time interval) blod pressure, cholesterol **Diffusion process**
- Values change a lot during a small time interval (start and stop smoking 0->1, number of visits before a given time point 0,1,2,...) **Counting process**
- The longer the follow-up time the more likely are changes
- lognitudinal data: registry based follow-up, repeated measures, e-equipments (smart tech.)

Heart transplant data

Mortality of patients on the waiting list for the Stanford heart transplant program.

- start, stop: Entry and exit time
- Event: and status for this interval of time (0=Alive, 1=Death)
- age: age-48 years
- year: year of acceptance (in years after 1 Nov 1967)
- surgery: prior bypass surgery 1=yes
- transplant: received transplant 1=yes
- id: patient id

start	stop	event	age	year	surgery	transplant	id
0	50	1	-17.2	0.12	0	0	1
0	6	1	3.8	0.25	0	0	2
0	1	0	6.3	0.27	0	0	3

Heart transplant data

Cox model with constant treatment (transplant) effect

```r
ci.lin(coxph(Surv(start, stop, event) ~
    age + year + surgery + transplant,
data=heart),Exp=TRUE)[,c(5:7)]
```

```
            exp(Est.)      2.5%     97.5%
age         1.0275390 1.0002875 1.0555330
year        0.8638585 0.7524197 0.9918021
surgery     0.5287657 0.2574423 1.0860419
transplant1 0.9898016 0.5351550 1.8306980
```

Transplant has no effect on mortality RR=0.99 !!!!!

How the risk is modified by time

- Change in the covariate value changes the risk instantly
- time lag of t $>$ 0 months,
- risk increases gradually until reaches some higher level or
- risk increases at first but decreases later. In this case the risk depends on the time when the covariate value changed.
- With diffuse processes change times can be defined in many ways! covariate value exceeds a certain level or covariate value increases by a certain amount during a (short) time interval.

HT data - the effect of transplantation

- Divide the follow-up time into e.g. at most four parts based on
- transplantation + 0 days
- transplantation + 30 days
- transplantation + 60 days.

Model a RR parameter for each interval.

Natural reference category can be interval before transplantation (set RR=1).

Time dependent effect of transplant -Lexis object

```
heart.Lx <-
Lexis(entry=list(age.time=age,
                 fu.time=0,
tx.time=decimal_date(tx.date)-decimal_date(accept.dt)),
duration=futime/365,
exit.status=as.character(fustat),
data=subset(jasa, is.na(tx.date) | fu.date != tx.date))
```

```
NOTE: entry.status has been set to "0" for all.
```

```
Warning in Lexis(entry = list(age.time = age, fu.time = 0,
```

Time dependent effect of transplant -Lexis object

```
heart.Lx <- within(heart.Lx, {cut.0 <- tx.time;
                              cut.1 <- tx.time+0.2;
                              cut.2 <- tx.time+0.4});
```

```
   age.time fu.time   tx.time   lex.dur lex.Cst lex.Xst
67 19.55099       0 0.1530055 0.7780822       0       1

   lex.id   birth.dt  accept.dt    tx.date    fu.date fusta
67     66 1952-09-03 1972-03-23 1972-05-18 1973-01-01

        age futime wait.time transplant mismatch hla.a2 msc
67 19.55099    284        56          1        3      0   1

   mscore reject     cut.2     cut.1     cut.0
67   1.02      0 0.5530055 0.3530055 0.1530055
```

Time dependent effect of transplant

Package Epi contains function mcutLexis, which splits the follow-up interval w.r.t. variables cut.0, cut.1 and cut.2: accept.dt acceptance == 0 tx.time transplant date (cut.0) tx.time +0.2 (cut.1) tx.time+0.4 (cut.2)

```
    age.time   fu.time   tx.time   lex.dur lex.Cst lex.Xst
156 19.55099 0.0000000 0.1530055 0.1530055       0       a
157 19.70400 0.1530055 0.3060109 0.2000000       a     a-b
158 19.90400 0.3530055 0.5060109 0.2000000     a-b   a-b-c
159 20.10400 0.5530055 0.7060109 0.2250767   a-b-c       1
```

```
    mscore reject     cut.2     cut.1     cut.0
156   1.02      0 0.5530055 0.3530055 0.1530055
157   1.02      0 0.5530055 0.3530055 0.1530055
158   1.02      0 0.5530055 0.3530055 0.1530055
159   1.02      0 0.5530055 0.3530055 0.1530055
```

Time dependent effect of transplant

Reference group is the time between acceptance and transplant operation, with RR=1

- a [transplant+0.2]
- a-b [transplant+0.2,transplant+0.4]
- a-b-c [transplant+0.4,]

```
print(format(as.data.frame(ci.exp(coxph(
      Surv(fu.time, fu.time + lex.dur, lex.Xst==1)
      ~ factor(lex.Cst),data=heart.Lx.cut))),digits=3))
```

```
                       exp(Est.)   2.5% 97.5%
factor(lex.Cst)a           1.371 0.7355  2.56
factor(lex.Cst)a-b         0.252 0.0501  1.26
factor(lex.Cst)a-b-c       0.528 0.1817  1.53
```

What should I consider when interpreting results

Main assumption has been that changes in the covariate value does not depend on outcome process. E.g. an intermediator:

- Treatment for HIV begins when CD4 has decreased below 350.
- Risk to contract AIDS (CD4$<$200) is higher for patients who have started treatment than patients who have not (CD4$>$350).
- A simple analysis where timedependent variable is defined by the time when CD4 reaches 350 may result in conclusion that the HIV treatment increases the risk of AIDS!
- In reality the risk would be higher without the treatment.
- More elaborate methods and adjustments are needed () consult a statistician).

How to proceed?

- Timedependent covariates can provide valuable information on treatment or risk factor effects.
- But: application of timedependent covariates in predicting e.g. survival can be challenging.
- Association of a timedependent covariate and the outcome can be instant, lagged or cumulative) selection of appropriate functional form?
- What causes the change of the covariate value?
- If the changetime of the covariate is completely random, the methods presented here are appropriate (e.g. RCT, heart transplantation example).