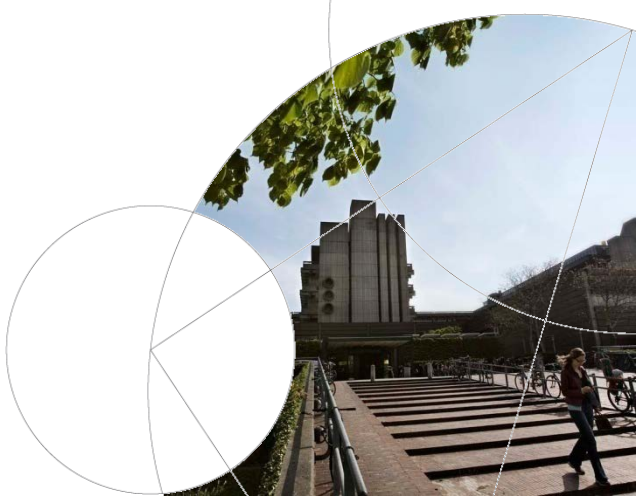
Survival analysis – pitfalls and more thoughts on time

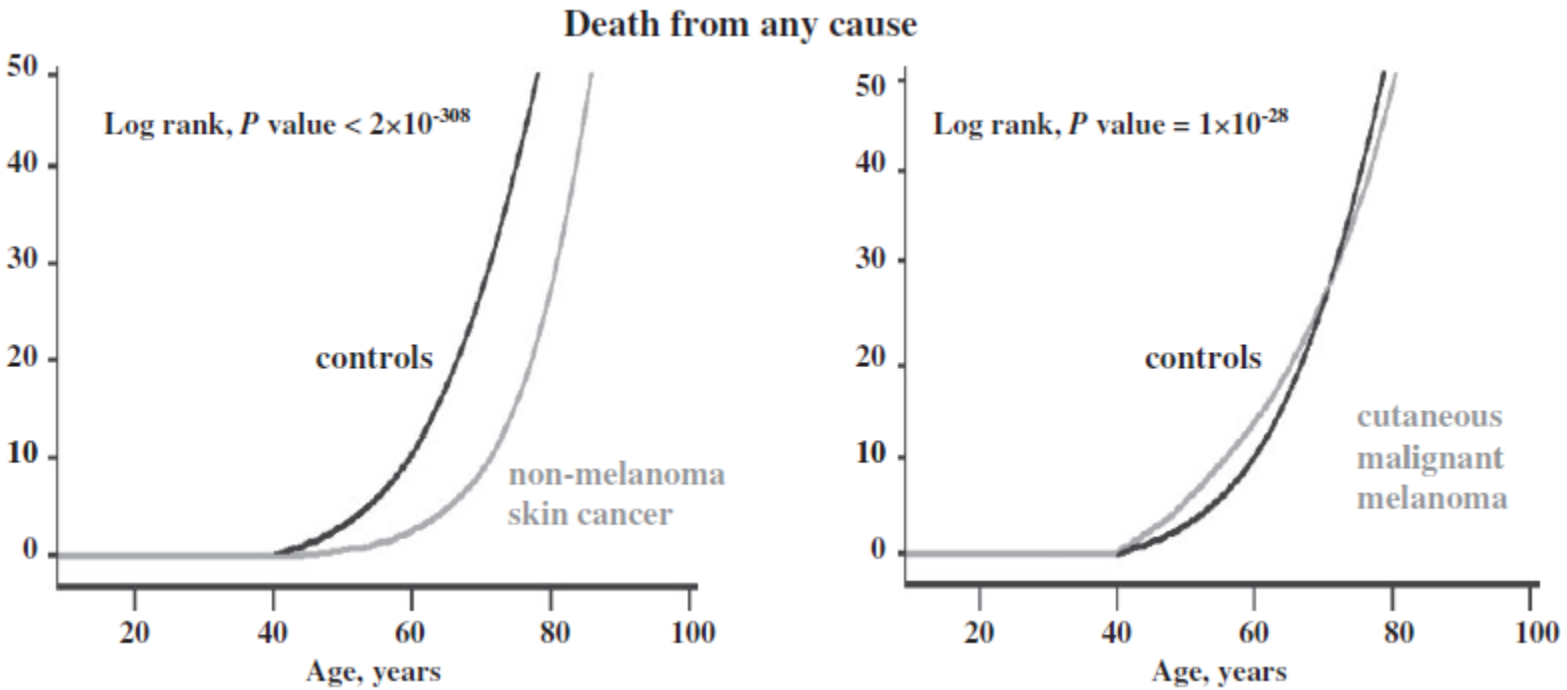
Theis Lange
Department of Biostatistics, University of Copenhagen
&
Center for Statistical Science, Peking University.

Mail: thlan@sund.ku.dk



Too good to be true?

The (1 minus) survival curves below compare a healthy control group with a non-melanoma skin cancer group.



Was this what you had expected?

Here is the method section of the paper

- We conducted a study of the entire Danish population above age 40 years from 1 January 1980 through 31 December 2006, comprising 4 412 568 individuals.
- Diagnoses and dates of skin cancer were drawn from the national Danish Cancer Registry, which identifies 98% of cancer cases in Denmark… All individuals with a diagnosis of non-melanoma skin cancer …from 1 January 1980 through 31 December 2006 were identified.
- Information on death from any cause was drawn from the national Danish Civil Registration System.
- We assessed the association between diagnoses of non-melanoma skin cancer … and death from any cause.
- The Cox regression models … and individuals were censored at event, death, permanent emigration or end of follow-up.

Take 5 min to discuss what was done with the person sitting next to you.

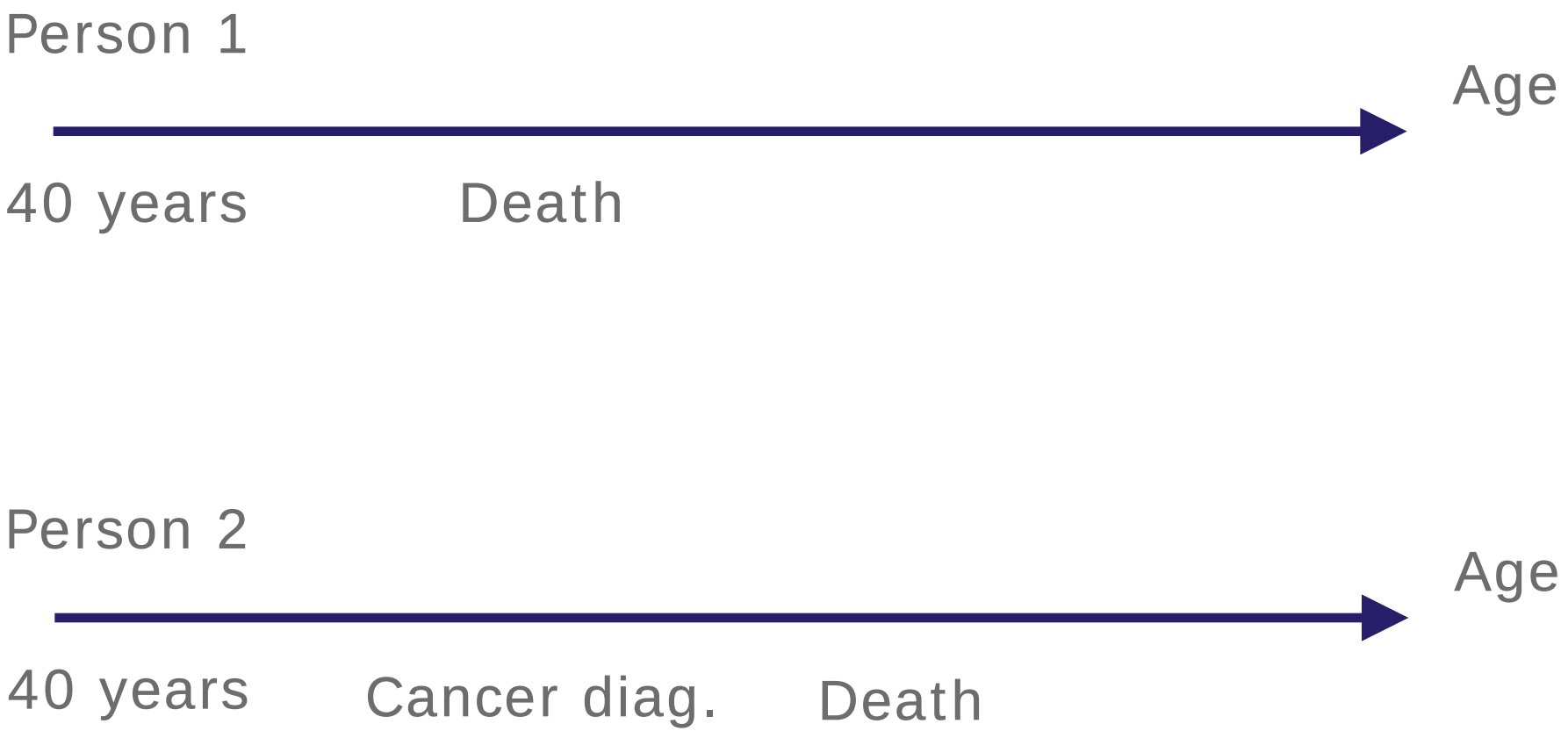
Int. J. Epidemiol. Advance Access published September 13, 2013

Published by Oxford University Press on behalf of the International Epidemiological Association

International Journal of Epidemiology 2013;1–11
doi:10.1093/ije/dyt168

Skin cancer as a marker of sun exposure associates with myocardial infarction, hip fracture and death from any cause

Peter Brøndum-Jacobsen,[1,3] Børge G Nordestgaard,[1,3] Sune F Nielsen[1] and Marianne Benn[2,3]*


Person 1
Age
40 years
Death
Person 2
Age
40 years
Cancer diag.
Death

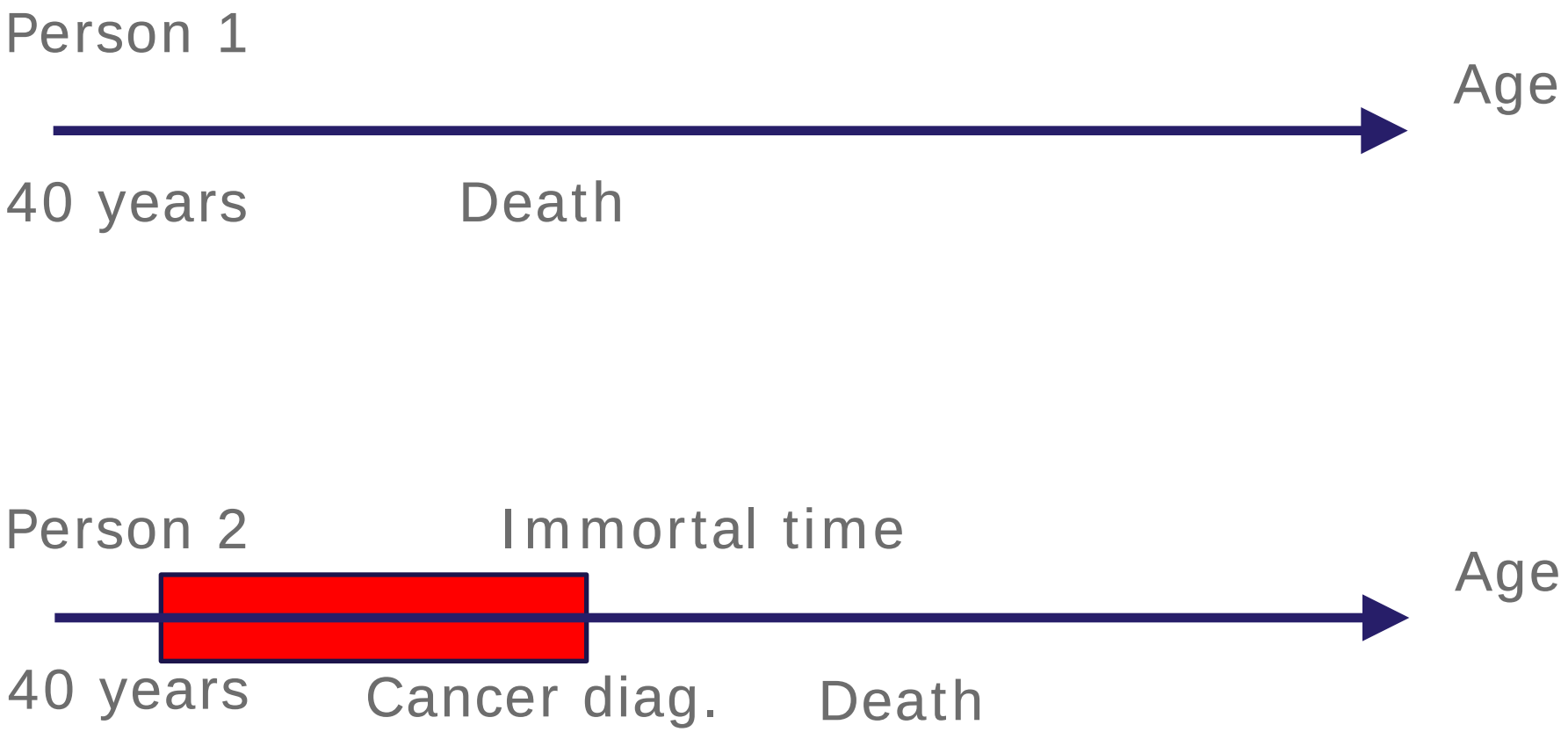
Our respond

Letters to the Editor

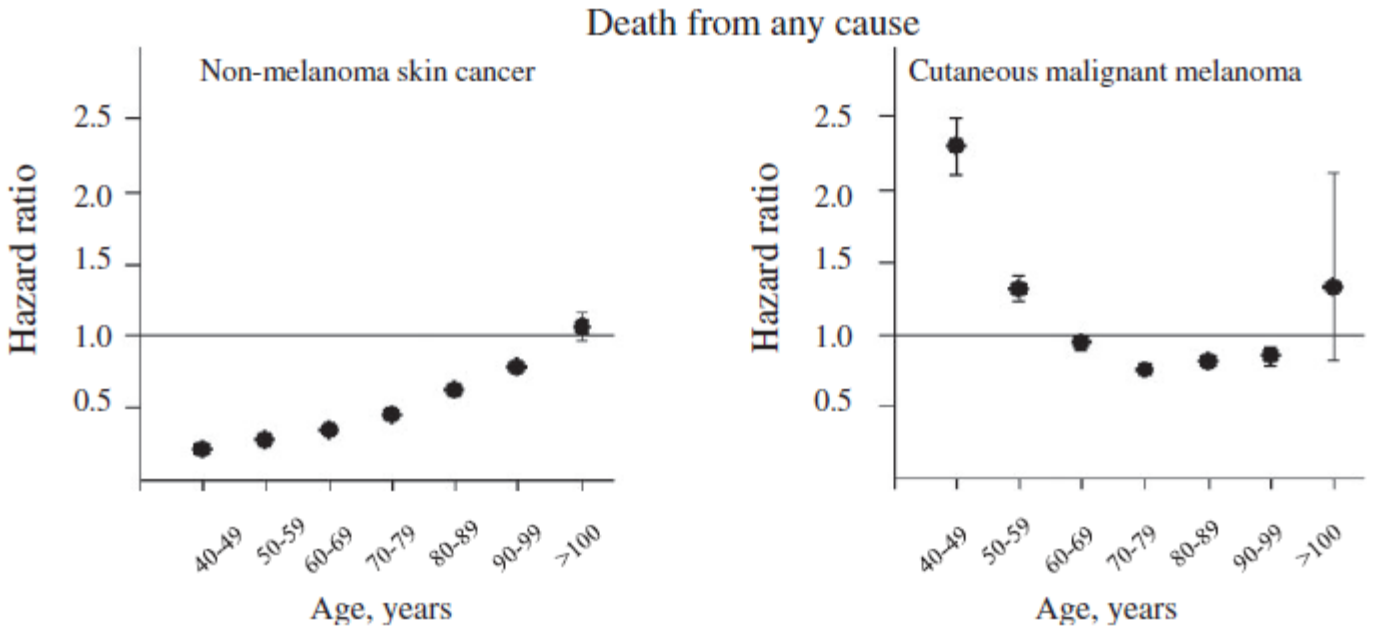
Skin cancer as a marker of sun exposure: a case of serious immortality bias

From Theis Lange* and Niels Keiding

Editorial in same issue where editor appoligized for mistakes in peer review process.


Person 1
Age
40 years
Death
Person 2
Immortal time
Age
40 years
Cancer diag.
Death

The effect of age stratification on immortal time bias



Exercise: Write two sentences explaining why you see the above.

The choice of time scale

A study is conducted over calendar time but the natural time variable may be time since treatment, e.g., the melanoma study.

Cohort studies are often conducted by recruiting a random sample of the population at the start of the study and then these subjects are followed for a number of years.

A natural time variable may be age rather than time on study which most often is an artificial time variable constructed by the investigators.

Vaccinations in Guinea-Bissau 1990-96

Rural Guinea-Bissau: 5274 children under 7 months of age visited two times at home, with an interval of six months. Information about vaccination (BCG, DTP, measles vaccine) collected at each visit and at second visit death during follow-up is registered. Some children were censored because they moved away during follow-up or survived until next visit.

Below are some of the variable names from the Bissau data.

`fuptime`	Follow-up time in days
`dead`	0 = censored, 1 = dead
`bcg`	1 = Yes, 2 = No
`agem`	Age at first visit in months

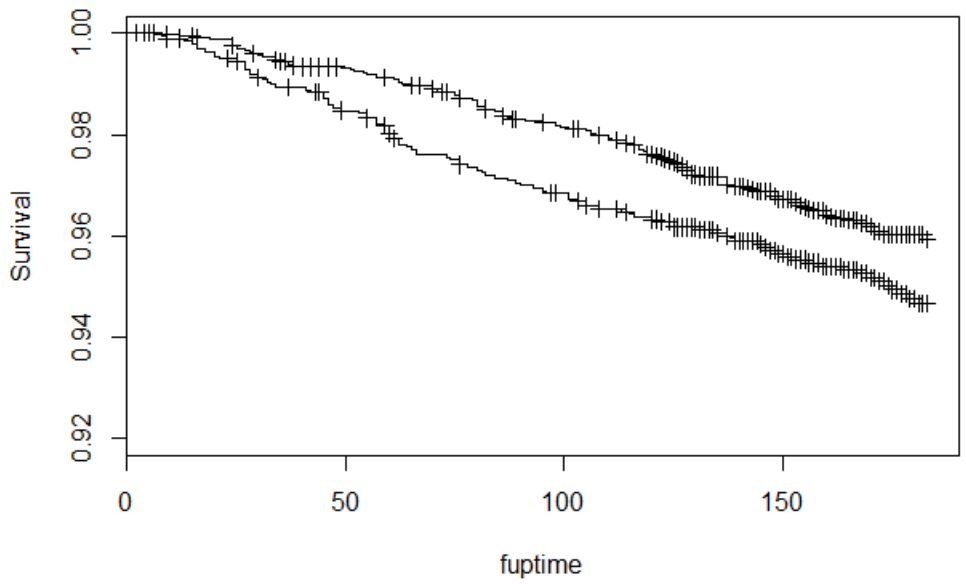
Is the risk of dying associated with vaccination?

Exposure	Outcome		
	Died	Survived	Total
BCG vaccinated	125 (3.8%)	3176	3301
not BCG vaccinated	97 (4.9%)	1876	1973
Total	222 (4.2%)	5052	5274

```
> chisq.test(bissauData$bcg, bissauData$dead)

        Pearson's Chi-squared test with Yates' continuity correction

data:  bissauData$bcg and bissauData$dead
X-squared = 3.6331, df = 1, p-value = 0.05664
```

Is this analysis OK?

Is the risk of dying associated with vaccination?

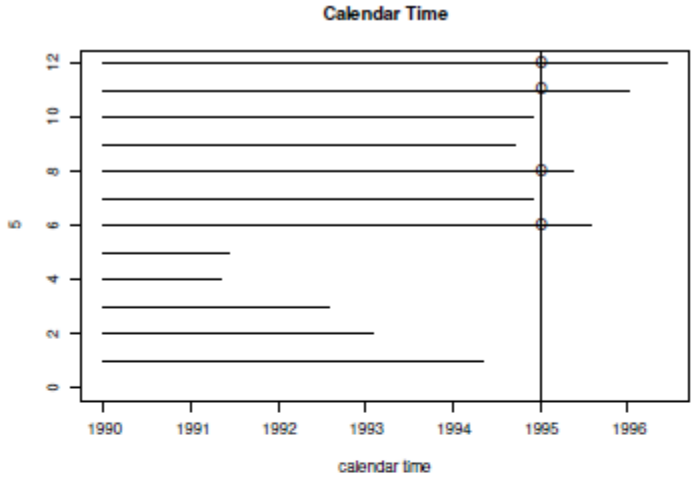


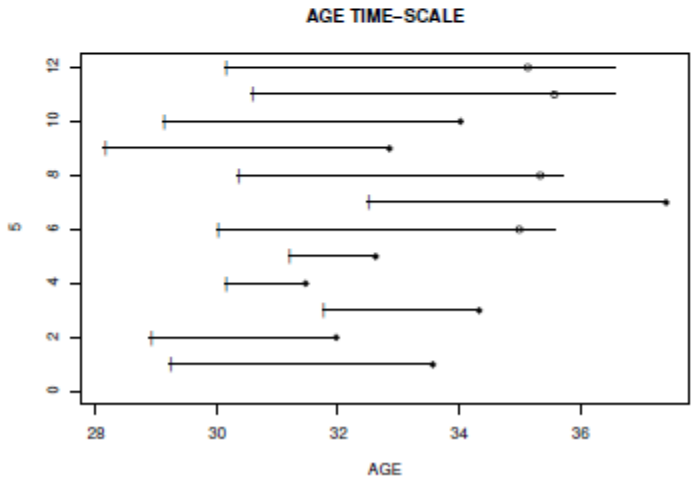

```
> plot(survfit(Surv(fuptime, dead)~bcg, data=bissauData), ylim=c(0.92,1), xlab="fuptime", ylab="Survival")
> survdiff(Surv(fuptime, dead)~bcg, data=bissauData)
Call:
survdiff(formula = Surv(fuptime, dead) ~ bcg, data = bissauData)

        N Observed Expected (O-E)^2/E (O-E)^2/V
bcg=1 3301      125      140      1.62      4.38
bcg=2 1973       97       82      2.76      4.38

 Chisq= 4.4  on 1 degrees of freedom, p= 0.0364
```

Is the risk of dying associated with vaccination?

```
> fit1Cox <- coxph(Surv(fuptime, dead)~bcg, data=bissauData)
> summary(fit1Cox)
Call:
coxph(formula = Surv(fuptime, dead) ~ bcg, data = bissauData)

  n= 5274, number of events= 222

      coef exp(coef) se(coef)     z Pr(>|z|)
bcg 0.2821    1.3260   0.1353 2.085   0.0371 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

    exp(coef) exp(-coef) lower .95 upper .95
bcg     1.326     0.7542     1.017     1.729

Concordance= 0.535  (se = 0.016 )
Rsquare= 0.001   (max possible= 0.51 )
Likelihood ratio test= 4.28  on 1 df,   p=0.03851
Wald test            = 4.35  on 1 df,   p=0.03707
Score (logrank) test = 4.38  on 1 df,   p=0.03645

>
```

Age as time variable: Delayed entry





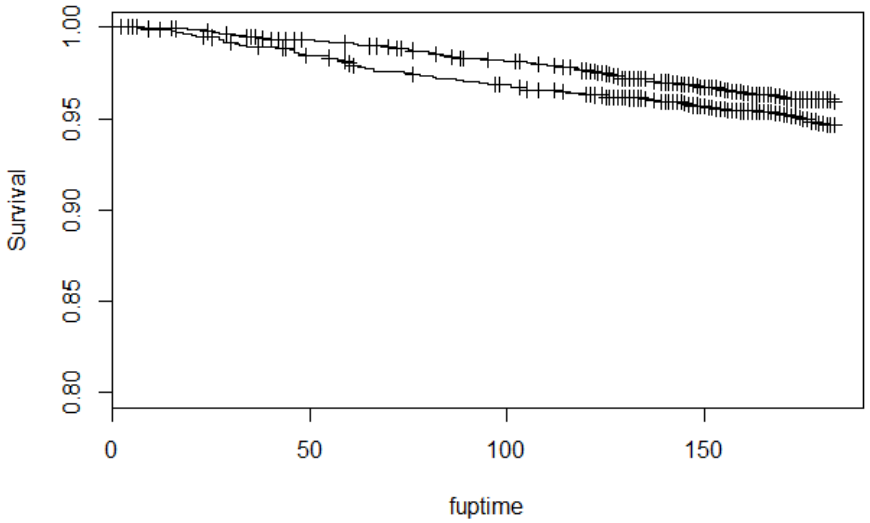
Age as time variable: Delayed entry

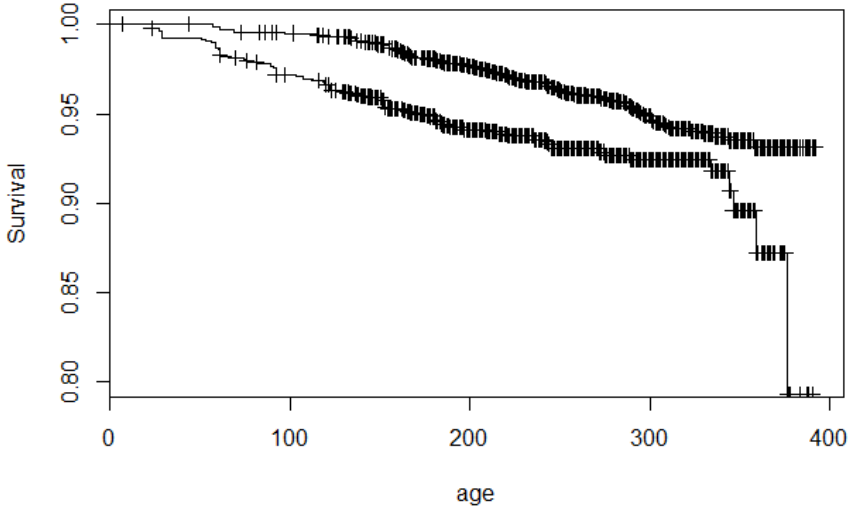
Subjects are only at risk at age of entry and onwards. They are not at risk before age of entry in our "World of analysis"!

Handling delayed entry is "easily" done by careful control of the RISK SET $R(t_i)$ at death time t_i.

```
> bissauData$outage <- bissauData$age+bissauData$fuptime
> fit2Cox <- coxph(Surv(age, outage, dead)~bcg, data=bissauData)
> summary(fit2Cox)
Call:
coxph(formula = Surv(age, outage, dead) ~ bcg, data = bissauData)

  n= 5274, number of events= 222

      coef exp(coef) se(coef)     z Pr(>|z|)
bcg 0.3552    1.4264   0.1407 2.525   0.0116 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

    exp(coef) exp(-coef) lower .95 upper .95
bcg     1.426     0.7011     1.083     1.879

Concordance= 0.533  (se = 0.017 )
Rsquare= 0.001   (max possible= 0.487 )
Likelihood ratio test= 6.26  on 1 df,   p=0.01233
Wald test            = 6.38  on 1 df,   p=0.01157
Score (logrank) test = 6.43  on 1 df,   p=0.01122
```

Age as time variable: Comparing the survival curves






```
plot(survfit(Surv(fuptime, dead)~bcg, data=bissauData), ylim=c(0.8,1),
     xlab="fuptime", ylab="Survival")
plot(survfit(Surv(age, outage, dead)~bcg, data=bissauData), ylim=c(0.8,1),
     xlab="age", ylab="Survival")
```

Time dependent explanatory variables

The Cox model can be expanded to include time-varying covariates

$$\lambda_i(t) = \lambda_0(t) \exp(\beta X_i(t)).$$

If the death times are $t_1, \ldots, t_d$ then it turns out that we "just" need to know the value of the covariates at the death times:

$$X_i(t_1), X_i(t_2), \ldots, X_i(t_d).$$

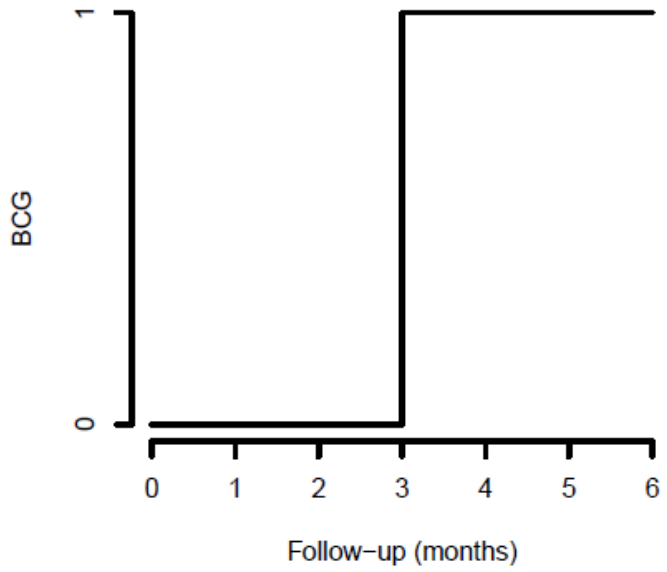
The covariate values at any time different from a death time are not used in the likelihood function.

The most simple time-varying covariate is a binary variable that is allowed to change once during follow-up, e.g. new BCG vaccinations registered between visits in the Bissau data:

$$X_i(t) = \begin{cases} 0 & \text{if no BCG before time } t \\ 1 & \text{if BCG-time} \leq t \end{cases}$$

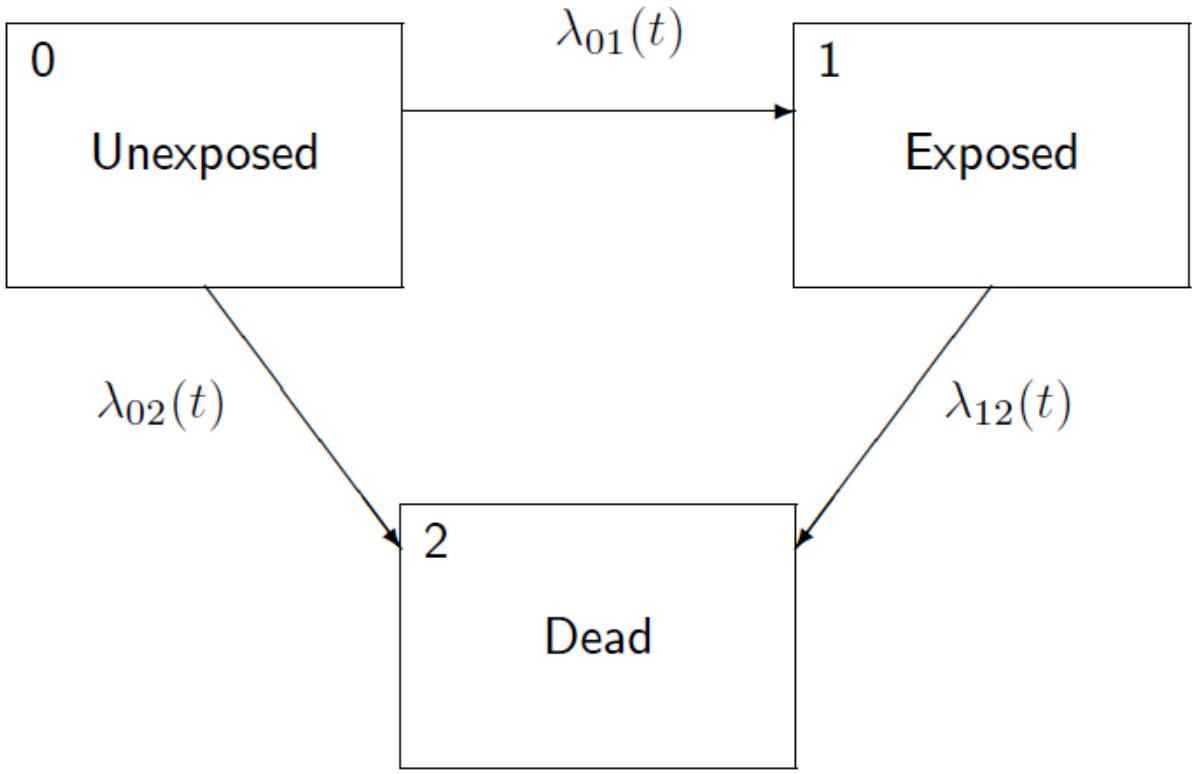
Time dependent explanatory variables

A child being BCG-vaccinated after 3 months of follow-up.



The time-varying covariate is 0 in the time interval 0 to 3 months and 1 for the rest of follow-up. For a child who was BCG vaccinated before first visit the time-varying covariate is one during all the follow-up.

Multi-state model



We want to compare $\lambda_{02}(t)$ and $\lambda_{12}(t)$. The transition rate $\lambda_{01}(t)$ is not modeled here.

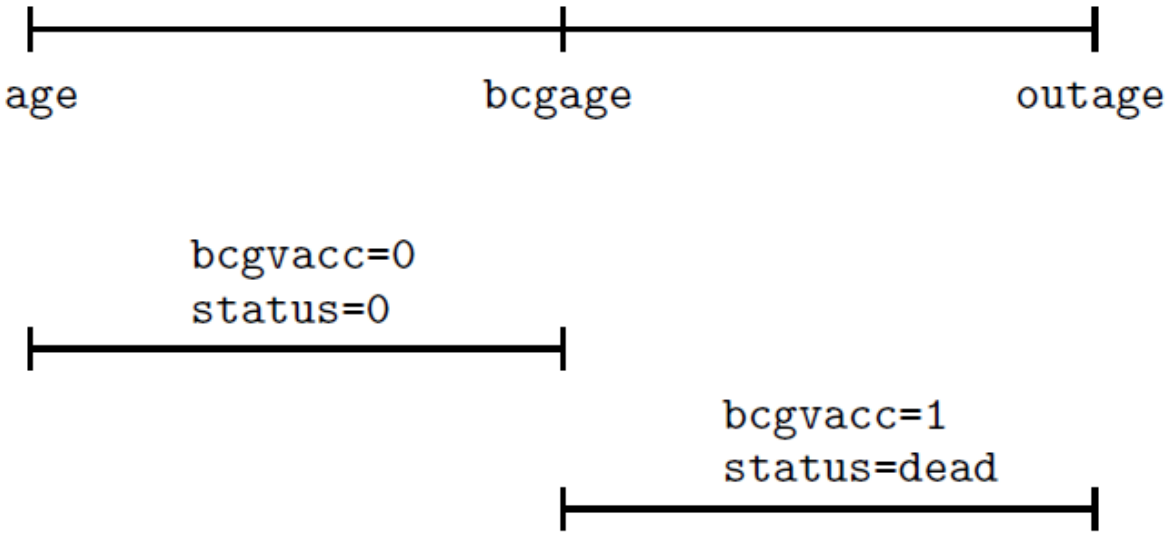
Time dependent explanatory variables

Instead of time of follow-up we will use age as time variable to illustrate the use of BCG as a time-varying covariate in the Bissau data. At visit 2 the vaccination cards were seen for the children at home and an age of BCG vaccination (bcgage) was calculated:

id	fuptime	dead	age	bcg	bcgage	outage
...						
486	159	0	199	1	107	358
487	183	0	97	1	20	280
488	183	0	43	2	174	226
489	137	1	140	1	40	277
490	183	0	165	1	46	348
...						
499	157	0	186	1	64	343
500	25	1	191	2	.	216
501	157	0	183	1	61	340
...						

Binary time-varying covariate in R

Splitting up persons with a changing time-varying covariate into two records:



and use delayed entry.

Thus, we need to generate a new data set.

Binary time-varying covariate in R

```
bissauData$bcgage[is.na(bissauData$bcgage)] <- 9999 # because NA can be
   bad when subsetting

# child not vac before outage or not at all
bissauDataTemp1 <- subset(bissauData, bcgage==9999 |  bcgage>outage)
bissauDataTemp1$bcgvacc  <- 0
bissauDataTemp1$entryage <- bissauDataTemp1$age
bissauDataTemp1$exitage  <- bissauDataTemp1$outage
bissauDataTemp1$status   <- bissauDataTemp1$dead



# child vac before initial age
bissauDataTemp2 <- subset(bissauData, bcgage<age)
bissauDataTemp2$bcgvacc  <- 1
bissauDataTemp2$entryage <- bissauDataTemp2$age
bissauDataTemp2$exitage  <- bissauDataTemp2$outage
bissauDataTemp2$status   <- bissauDataTemp2$dead
```


Binary time-varying covariate in R

```
# child vac after initial age, but before end of follow-up
bissauDataTemp3 <- subset(bissauData, (age<bcgage) & (bcgage<outage))
bissauDataTemp4 <- subset(bissauData, (age<bcgage) & (bcgage<outage))
bissauDataTemp3$bcgvacc  <- 0
bissauDataTemp4$bcgvacc  <- 1
bissauDataTemp3$entryage <- bissauDataTemp3$age
bissauDataTemp4$entryage <- bissauDataTemp4$bcgage
bissauDataTemp3$exitage  <- bissauDataTemp3$bcgage
bissauDataTemp4$exitage  <- bissauDataTemp4$outage
bissauDataTemp3$status   <- 0
bissauDataTemp4$status   <- bissauDataTemp4$dead

# collect all
bissauDataSplit <- rbind(bissauDataTemp1, bissauDataTemp2,
   bissauDataTemp3, bissauDataTemp4)
bissauDataSplit <- bissauDataSplit[order(bissauDataSplit$id), ]
```

Binary time-varying covariate in R

id	fuptime	dead	age	bcg	bcgage	outage	bcgvacc	entryage	exitage	status
488	183	0	43	2	174	226	0	43	174	0
488	183	0	43	2	174	226	1	174	226	0

```
fit3Cox <- coxph(Surv(entryage, exitage, status)~bcgvacc,
        data=bissauDataSplit)
summary(fit3Cox)
```

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
bcgvacc	1	-1.08278	0.14046	59.4286	<.0001	0.339	0.257	0.446

Note 1: No need to care for the repeated rows for some persons – why?
Note 2: Bigger effect size (old analysis HR=1.4; now HR=2.9) – why?

Stanford heart transplant data (p. 235)

In an article (Crowley and Hu, *J Amer. Statist Assoc.* 1977) on the Stanford Heart Transplantation Study, patients identified as being eligible (N=103) for a heart transplant were followed until death or censorship. In total, 65 received a transplant during follow-up, whereas 38 did not.

The purpose is to assess whether transplanted patients survived longer and at the exercises we will do some of the analyses.

Stanford heart transplant data (p. 235)

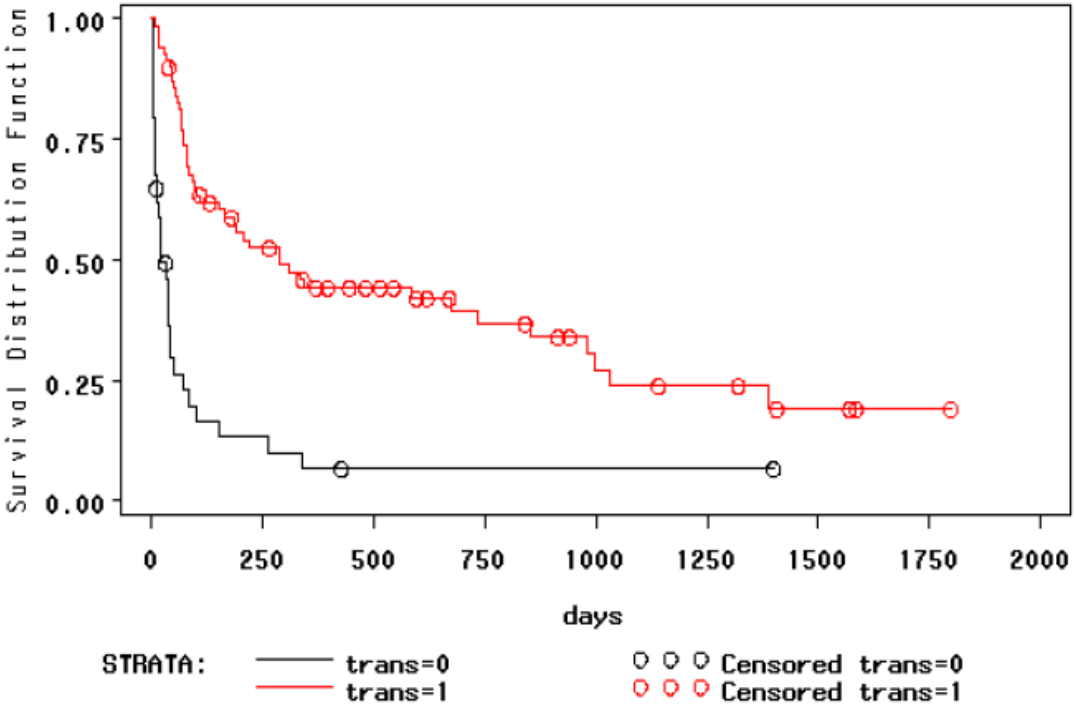
`age`	age (in years) at entry into the study.
`cens`	0 = Censoring
	1 = Dead
`days`	number of days from entry to dead/censoring.
`trans`	1 = if the person had a heart transplantation
	0 = otherwise.
`wait`	number of days from entry to transplantation
	NB: if trans = 0 then wait = -1
`mismatch`	1 = mismatch between HLA type in donor and patient
	0 = no mismatch
	NB: if trans = 0 then mismatch = -1.

Stanford heart transplant data (p. 235)

Obs	age	cens	days	trans	wait	mismatch
52	56	1	90	1	27	1
53	53	1	96	1	67	0
54	48	1	100	1	46	0
55	41	1	102	0	-1	-1
56	28	0	109	1	96	1
57	46	1	110	1	60	0
58	23	0	131	1	21	1
59	41	1	149	0	-1	-1
60	47	1	153	1	26	0
61	43	1	165	1	4	0
62	26	0	180	1	13	0
63	52	1	186	1	160	1
64	47	1	188	1	41	0
65	51	1	207	1	139	1
66	51	1	219	1	83	1
67	8	1	263	0	-1	-1
68	47	0	265	1	28	0
69	48	1	285	1	32	1
70	19	1	285	1	57	0
71	49	1	308	1	28	0

Stanford heart transplant data (p. 235)

Why are they wrong?!

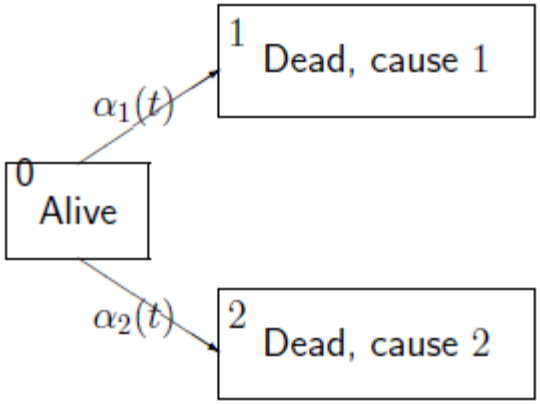


COMPETING RISKS

Competing risks

- In studies of all-cause mortality, rates (hazards) can be computed from risks (probabilities, cumulative incidences) and vice versa - in other words the two functions contain *equivalent* information
- In studies of events which will not eventually happen for every one in the population, this is no longer the case and death (and maybe other events) are *competing risks* which need to be addressed
- In such cases, the risk of a given cause depends on the rates for *all* competing causes
- Therefore, using '1-Kaplan-Meier for a single cause' as a risk estimator is (upward) biased
- The magnitude of the bias depends on the frequency of the competing events

Competing risks



For simplicity we consider two causes of death, and let interest focus on learning about the hazard rates $\alpha_1(t)$ and $\alpha_2(t)$. These gives the cause-specific chance of dying of cause k if alive at time t. We may integrate the cause specific hazards

$$\Gamma_k(t) = \int_0^t \alpha_k(s)ds \; k = 1, 2$$

The total mortality is given by

$$\begin{aligned} \lambda(t) &= \alpha_1(t) + \alpha_2(t) \\ \Lambda(t) &= \Gamma_1(t) + \Gamma_2(t) \end{aligned}$$

and the survival function is given as

$$S(t) = \exp(-\Lambda(t))$$

where $\Lambda(t) = \int_0^t \lambda(s)ds$.

Competing risks

- Cause specific hazard (and hazard ratios) can be estimated simply by censoring on all other events and then use coxph-function.

- However, survival functions (Kaplan-Meier) are biased.

- Instead cumulative incidence curves can be computed (big topic we will not get into).

Competing risks

- In studies of all-cause mortality, rates (hazards) can be computed from risks (probabilities, cumulative incidences) and vice versa - in other words the two functions contain *equivalent* information
- In studies of events which will not eventually happen for every one in the population, this is no longer the case and death (and maybe other events) are *competing risks* which need to be addressed
- In such cases, the risk of a given cause depends on the rates for *all* competing causes
- Therefore, using '1-Kaplan-Meier for a single cause' as a risk estimator is (upward) biased
- The magnitude of the bias depends on the frequency of the competing events

ALTERNATIVE MODELS FOR SURVIVAL DATA

The additive hazard models

ORIGINAL ARTICLE

Additive Interaction in Survival Analysis

Use of the Additive Hazards Model

Naja Hulvej Rod,[a] *Theis Lange,*[b] *Ingelise Andersen,*[a] *Jacob Louis Marott,*[c] *and Finn Diderichsen*[a]

(*Epidemiology* 2012;23: 733–737)

The additive hazard models

- The hazard ratio estimates from Cox can be hard to interpret.
- Effect measures capturing number of events directly might be more appropriate.
- The Aalen additive hazard model is a solution.
- From “birth” the model is non-parametric.
- With the covariates (E_1 and E_2) the hazard at time t is given by

$$\gamma_0(t) + \beta_1(t)E_1 + \beta_1(t)E_2 + \beta_3(t)(E_1 \times E_2)$$

- Estimated by OLS at each event time.

6S: The additive hazard models

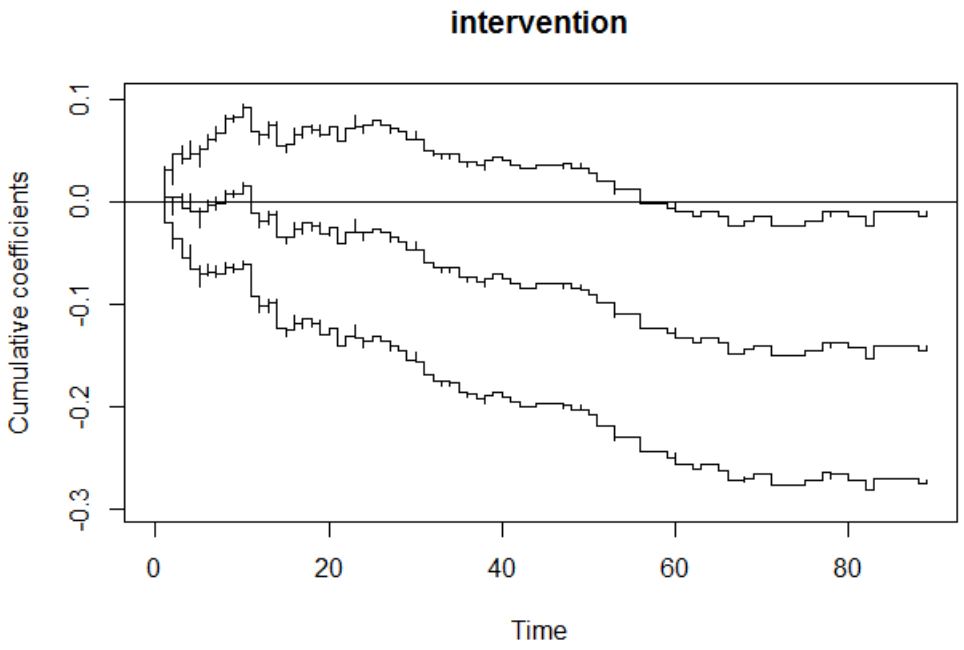
- In R additive hazard models can be estimated using the aalen function from the timereg package.

```
> library(timereg)
> fit1Aalen <- aalen(Surv(time_to_death, mortality_90days)~intervention, data=kidneyData)
> summary(fit1Aalen)
Additive Aalen Model

Test for nonparametric terms

Test for non-significant effects
             Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                          13.90                0.000
intervention                          2.34                0.252

Test for time invariant effects
                  Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                        0.2790                       0.000
intervention                       0.0437                       0.793
                    Cramer von Mises test p-value H_0:constant effect
(Intercept)                        3.1000                       0.000
intervention                       0.0345                       0.648



  Call:
aalen(formula = Surv(time_to_death, mortality_90days) ~ intervention,
    data = kidneyData)
```

- Why no parameter estimate?

6S: The additive hazard models

- You can only get plots of cumulated coefficient.
- That is plots of $\int_0^t \beta_1(s)\mathrm{d}s$ etc.
- In R: **`plots(fit1Aalen)`**



6S: The additive hazard models

- The aalen model also exists in a semi-parametric version.
- Here the hazard is modeled as:

$$\gamma_0(t) + \beta_1 E_1 + \beta_1 E_2 + \beta_3 (E_1 \times E_2)$$

```
> fit2Aalen <- aalen(Surv(time_to_death, mortality_90days)~const(intervention), data=kidneyData)
> summary(fit2Aalen)
Additive Aalen Model

Test for nonparametric terms

Test for non-significant effects
            Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                          16.3                    0

Test for time invariant effects
                 Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                       0.283                           0
                   Cramer von Mises test p-value H_0:constant effect
(Intercept)                       2.96                            0

Parametric terms :
                     Coef.    SE Robust SE    z P-val
const(intervention) -0.002 0.001     0.001 -1.8 0.072

  Call:
aalen(formula = Surv(time_to_death, mortality_90days) ~ const(intervention),
    data = kidneyData)
```

Cox vs. Aalen model

TABLE 2. Comparison of the Estimation of the Interaction Between Education and Smoking on Risk of Lung Cancer in the Cox Proportional Hazards Model and the Additive Hazards Model Based on Data on 73,145 Men and Women From the Danish Social Inequality in Cancer Database

	Cox Proportional Hazards Model Hazard Ratio[a] (95% CI)	Additive Hazards Model No. Additional Lung Cancer Cases per 10,000 Person Years[a] (95% CI)
Model output		
Short vs. long education	1.32 (1.03 to 1.71)	−0.5 (−1.9 to 1.0)
Smoker vs. non-smoker	7.15 (6.01 to 8.51)	24.2 (22.0 to 26.4)
Short education*smoker	1.18 (0.89 to 1.55)	18.5 (14.0 to 23.0)
Stratified analysis		
Long education		
Smoker vs. non-smoker	7.15 (6.01 to 8.51)	24.2 (22.0 to 26.4)
Short education		
Smoker vs. non-smoker	8.43 (6.79 to 10.5)	42.7 (38.9 to 46.5)
Test for interaction	$P = 0.24$	$P < 0.001$
Joint effects		
Long education, non-smoker[b]	1.00	1.00
Short education, non-smoker	1.32 (1.03 to 1.71)	−0.5 (−1.9 to 1.0)
Long education, smoker	7.15 (6.01 to 8.51)	24.2 (22.0 to 26.4)
Short education, smoker	11.16 (9.37 to 13.3)	42.2 (38.6 to 45.8)

[a]Adjusted for sex, age, and cohort.
[b]Reference category.
CI indicates confidence interval.

The exercise

There is a mandatory exercise to be handed in within two weeks.