The resurrection of time as a continuous concept in biostatistics, demography and epidemiology

Bendix Carstensen Steno Diabetes Center,
Gentofte, Denmark
& Department of Biostatistics, University of Copenhagen
bxc@steno.dk
http://BendixCarstensen.com

IMBEI, Mainz, Germany, 20 September 2016

http://BendixCarstensen.com/AdvCoh/talks/

Inference in Multistate models

P.K. Andersen & N. Keiding
Interpretability and Importance of Functionals in Competing Risks and Multistate Models, *Stat Med, 2011* [1]:

1. Do not condition on the future
2. Do not regard individuals at risk after they have died
3. Stick to this world

Conditioning on the future

- . . . also known as "Immortal time bias", see *e.g.* S. Suissa: Immortal time bias in pharmaco-epidemiology, *Am. J. Epidemiol*, 2008 [2].
- Including persons' follow-up in the wrong state
- . . . namely one reached some time in the future
- Normally caused by classification of **persons** instead of classification of **follow-up time**

Why these mistakes?

- Time is usually absent from survival analysis results
- . . . because time is taken to be a **response** variable observed for each **person**
- Unit of analysis is often seen as the person
- Non/Semi-parametric survival model interface invites this misconception
- **Persons** classified by exposure (the latest, often)
- The **real** unit of observation should be person-**time**
- . . . intervals of time, each with different **value** of
  - time
  - other covariates

Time

- Time is a **covariate** — determinant of rates
- **Response** variable in survival / follow-up is bivariate:
  - **Differences** on the timescale (**risk** time, "exposure")
  - **Events**
- The relevant unit of observation is person-time:
  - small intervals of follow-up — "empirical rates"
  - (d_{it}, y_{it}): (event, (sojourn) time) for individual i at time t.
  - y is the **response** time, t is the **covariate** time
- Covariates relate to each interval of follow-up
- Allows **multiple** timescales, *e.g.* age, duration, calendar time

"Stick to this world"

In the paper by Andersen & Keiding this is primarily aimed at the use of "net survival", that is the calculation of

$$\exp\left(-\int_0^t \lambda_c(s)\,\mathrm{d}s\right)$$

for a single cause of death
— formally for a non-exhaustive exit rate from a state.

Survival probability in the situation where:

1. all other causes of death are absent
2. the mortality, λ_c from cause c is unchanged

. . . which is indeed **not** of this world.

Sticking to this world

- A further feature of "this world":
- it is **continuous**
- no thresholds in the effect of time
- specifically, death and disease rates vary **smoothly** by
  - age
  - calendar time
  - disease duration
  - ...

A look at the Cox model

$$\lambda(t, x) = \lambda_0(t) \times \exp(x'\beta)$$

A model for the rate as a function of t and x.

The covariate t has a special status:

- Computationally, because all individuals contribute to (some of) the range of t.
- ...the scale along which time is split (the risk sets)
- Conceptually t is just a covariate that varies within individual.
- Cox's approach profiles $\lambda_0(t)$ out from the model

The Cox-likelihood as profile likelihood

- One parameter per death time to describe the effect of time (i.e. the chosen timescale).

$$\log\big(\lambda(t, x_i)\big) = \log\big(\lambda_0(t)\big) + \beta_1 x_{1i} + \cdots + \beta_p x_{pi} = \alpha_t + \eta_i$$

- Profile likelihood:
  - Derive estimates of α_t as function of data and βs — assuming constant rate between death times
  - Insert in likelihood, now only a function of data and βs
  - Turns out to be Cox's partial likelihood

The Cox-likelihood: mechanics of computing

- The likelihood is computed by suming over risk-sets:

$$\ell(\eta) = \sum_t \log\left(\frac{\mathrm{e}^{\eta_{\text{death}}}}{\sum_{i\in\mathcal{R}_t}\mathrm{e}^{\eta_i}}\right)$$

- this is essentially splitting follow-up time at event- (and censoring) times
- . . . repeatedly in every cycle of the iteration
- . . . simplified by not keeping track of risk time
- . . . but only works along **one** time scale

$$\log\big(\lambda(t, x_i)\big) = \log\big(\lambda_0(t)\big) + \beta_1 x_{1i} + \cdots + \beta_p x_{pi} = \alpha_t + \eta_i$$

- Suppose the time scale has been divided into small intervals with at most one death in each:
- Empirical rates: (d_{it}, y_{it}) — each t has at most one $d_{it} = 0$.
- Assume w.l.o.g. the ys in the empirical rates all are 1.
- Log-likelihood contributions that contain information on a specific time-scale parameter α_t will be from:
  - the (only) empirical rate $(1, 1)$ with the death at time t.
  - all other empirical rates $(0, 1)$ from those who were at risk at time t.

Note: There is one contribution from each person at risk to this part of the log-likelihood:

$$\begin{aligned}
\ell_t(\alpha_t, \beta) &= \sum_{i\in\mathcal{R}_t} d_i \log(\lambda_i(t)) - \lambda_i(t) y_i \\
&= \sum_{i\in\mathcal{R}_t} \left\{ d_i(\alpha_t + \eta_i) - \mathrm{e}^{\alpha_t+\eta_i} \right\} \\
&= \alpha_t + \eta_{\text{death}} - \mathrm{e}^{\alpha_t} \sum_{i\in\mathcal{R}_t} \mathrm{e}^{\eta_i}
\end{aligned}$$

where η_{death} is the linear predictor for the person that died.

The derivative w.r.t. α_t is:

$$\mathrm{D}_{\alpha_t}\ell_t(\alpha_t,\beta) = 1 - \mathrm{e}^{\alpha_t}\sum_{i\in\mathcal{R}_t}\mathrm{e}^{\eta_i} = 0 \quad \Leftrightarrow \quad \mathrm{e}^{\alpha_t} = \frac{1}{\sum_{i\in\mathcal{R}_t}\mathrm{e}^{\eta_i}}$$

If this estimate is fed back into the log-likelihood for α_t, we get the **profile likelihood** (with α_t "profiled out"):

$$\log\left(\frac{1}{\sum_{i\in\mathcal{R}_t}\mathrm{e}^{\eta_i}}\right) + \eta_{\text{death}} - 1 = \log\left(\frac{\mathrm{e}^{\eta_{\text{death}}}}{\sum_{i\in\mathcal{R}_t}\mathrm{e}^{\eta_i}}\right) - 1$$

which is the same as the contribution from time t to Cox's partial likelihood.

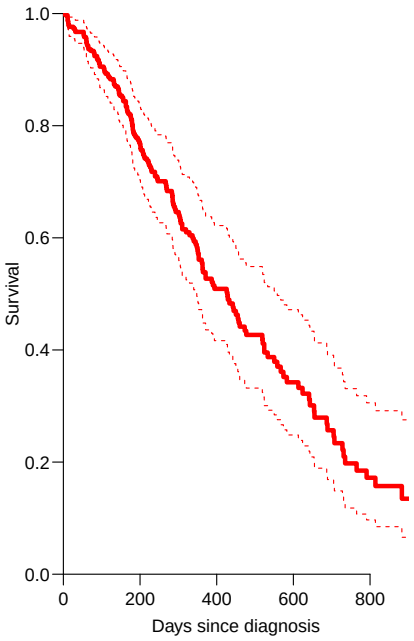
Splitting the dataset a priori

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — each individual contributes many empirical rates
- (one per risk-set contribution in Cox-modelling)
- From each empirical rate we get:
  - Poisson-response d
  - Risk time $y \to \log(y)$ as offset
  - Covariate value for the timescale (time since entry, current age, current date, ...)
  - other covariates
- Contributions not independent, but likelihood is a product
- Same likelihood as for independent Poisson variates
- Modelling is by standard `glm` Poisson

Example: Mayo Clinic lung cancer

- Survival after lung cancer
- Covariates:
  - Age at diagnosis
  - Sex
  - Time since diagnosis
- Cox model
- Split data:
  - Poisson model, time as factor
  - Poisson model, time as spline

Mayo Clinic lung cancer 60 year old woman



Example: Mayo Clinic lung cancer I

```
> library( survival )
> library( Epi )
> Lung <- Lexis( exit = list( tfe=time ),
+                exit.status = factor(status,labels=c("Alive","Dead")),
+                data = lung )

NOTE: entry.status has been set to "Alive" for all.
NOTE: entry is assumed to be 0 on the tfe timescale.
```


Example: Mayo Clinic lung cancer II

```
> mL.cox <- coxph( Surv( tfe, tfe+lex.dur, lex.Xst=="Dead" ) ~
+                  age + factor( sex ),
+                  method="breslow", eps=10^-8, iter.max=25, data=Lung )
> Lung.s <- splitLexis( Lung,
+                       breaks=c(0,sort(unique(Lung$time))),
+                       time.scale="tfe" )
> Lung.S <- splitLexis( Lung,
+                       breaks=c(0,sort(unique(Lung$time[Lung$lex.Xst=="Dead"]))),
+                       time.scale="tfe" )
> summary( Lung.s )

Transitions:
     To
From    Alive Dead  Records:  Events: Risk time:  Persons:
  Alive 19857  165     20022      165      69593       228

> summary( Lung.S )
```

Example: Mayo Clinic lung cancer III

```
Transitions:
     To
From    Alive Dead  Records:  Events: Risk time:  Persons:
  Alive 15916  165     16081      165      69593       228

> subset( Lung.s, lex.id==96 )[,1:11]

     lex.id tfe lex.dur lex.Cst lex.Xst inst time status age sex ph.ecog
9235     96   0       5   Alive   Alive   12   30      2  72   1       2
9236     96   5       6   Alive   Alive   12   30      2  72   1       2
9237     96  11       1   Alive   Alive   12   30      2  72   1       2
9238     96  12       1   Alive   Alive   12   30      2  72   1       2
9239     96  13       2   Alive   Alive   12   30      2  72   1       2
9240     96  15      11   Alive   Alive   12   30      2  72   1       2
9241     96  26       4   Alive    Dead   12   30      2  72   1       2

> nlevels( factor( Lung.s$tfe ) )

[1] 186
```


Example: Mayo Clinic lung cancer IV

```
> system.time(
+ mLs.pois.fc <- glm( lex.Xst=="Dead" ~ - 1 + factor( tfe ) +
+                                age + factor( sex ),
+                                offset = log(lex.dur),
+                     family=poisson, data=Lung.s, eps=10^-8, maxit=25 )
+             )

   user  system elapsed
 10.828   0.012  10.837

> length( coef(mLs.pois.fc) )

[1] 188

> system.time(
+ mLS.pois.fc <- glm( lex.Xst=="Dead" ~ - 1 + factor( tfe ) +
+                                age + factor( sex ),
+                                offset = log(lex.dur),
+                     family=poisson, data=Lung.S, eps=10^-8, maxit=25 )
+             )
```


Example: Mayo Clinic lung cancer V

```
   user  system elapsed
  3.258   0.000   3.257

> length( coef(mLS.pois.fc) )

[1] 142

> t.kn <- c(0,25,100,500,1000)
> dim( Ns(Lung.s$tfe,knots=t.kn) )

[1] 20022     4

> system.time(
+ mLs.pois.sp <- glm( lex.Xst=="Dead" ~ Ns( tfe, knots=t.kn ) +
+                                age + factor( sex ),
+                     offset = log(lex.dur),
+                     family=poisson, data=Lung.s, eps=10^-8, maxit=25 )
+             )
```

Example: Mayo Clinic lung cancer VI

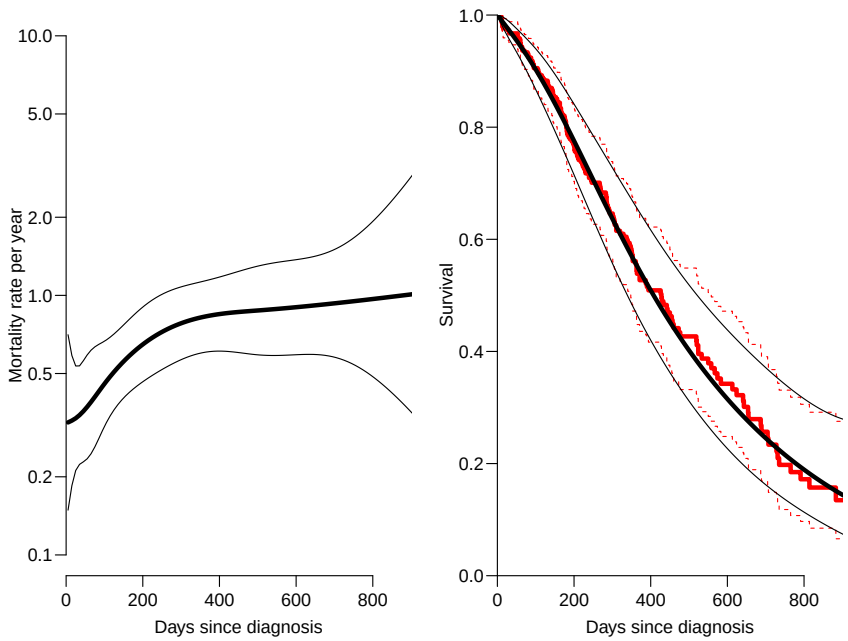
```
   user  system elapsed
  0.173   0.000   0.172
```

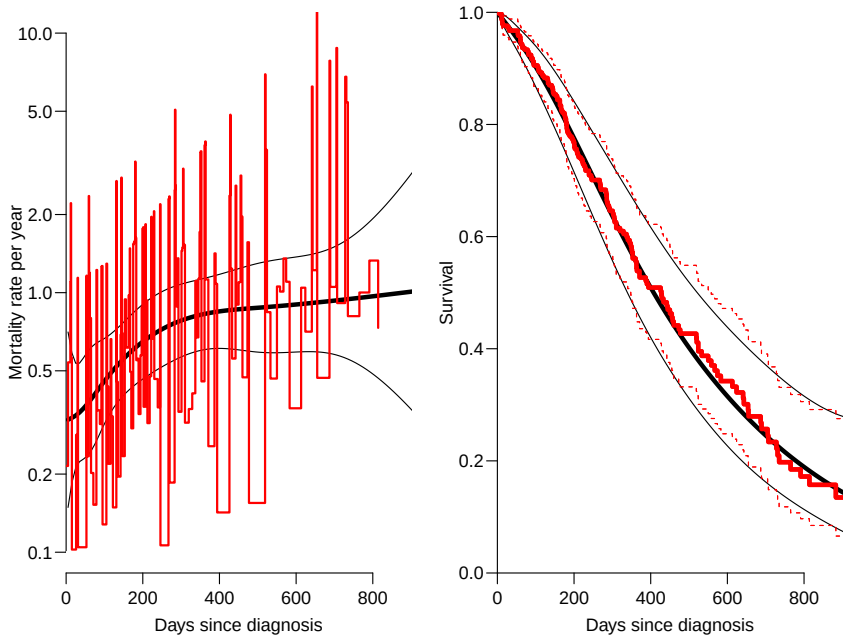
```
> ests <-
+ rbind( ci.exp(mL.cox),
+        ci.exp(mLs.pois.fc,subset=c("age","sex")),
+        ci.exp(mLS.pois.fc,subset=c("age","sex")),
+        ci.exp(mLs.pois.sp,subset=c("age","sex")) )
> cmp <- cbind( ests[c(1,3,5,7)  ,],
+               ests[c(1,3,5,7)+1,] )
> rownames( cmp ) <- c("Cox","Poisson-factor","Poisson-factor (D)","Poisson-spline
> colnames( cmp )[c(1,4)] <- c("age","sex")
```

```
> round( cmp, 7 )
```


Example: Mayo Clinic lung cancer VII

	age	2.5%	97.5%	sex	2.5%	97.5%
Cox	1.017158	0.9989388	1.035710	0.5989574	0.4313720	0.8316487
Poisson-factor	1.017158	0.9989388	1.035710	0.5989574	0.4313720	0.8316487
Poisson-factor (D)	1.017332	0.9991211	1.035874	0.5984794	0.4310150	0.8310094
Poisson-spline	1.016189	0.9980329	1.034676	0.5998287	0.4319932	0.8328707





Deriving the survival function

```
> mLs.pois.sp <- glm( lex.Xst=="Dead" ~ Ns( tfe, knots=t.kn ) +
+                                  age + factor( sex ),
+                       offset = log(lex.dur),
+                       family=poisson, data=Lung.s, eps=10^-8, maxit=25 )
```

```
> CM <- cbind( 1, Ns( seq(10,1000,10)-5, knots=t.kn ), 60, 1 )
> lambda <- ci.exp( mLs.pois.sp, ctr.mat=CM )
> Lambda <- ci.cum( mLs.pois.sp, ctr.mat=CM, intl=10 )[,-4]
> survP  <- exp(-rbind(0,Lambda))
```

Code and output avaialble in
`http://bendixcarstensen.com/AdvCoh/WNtCMa/`

What the Cox-model really is

Taking the life-table approach *ad absurdum* by:

- dividing time very finely and
- modeling one covariate, the time-scale, with one parameter per distinct value.
- the **model** for the time scale is really with exchangeable time-intervals.
- $\Rightarrow$ difficult to access the baseline hazard.
- $\Rightarrow$ uninitiated tempted to show survival curves where irrelevant

Models of this world

- Replace the α_ts by a parametric function $f(t)$ with a limited number of parameters, for example:
  - Piecewise constant
  - Splines (linear, quadratic or cubic)
  - Fractional polynomials
- Brings model into "this world":
  - smoothly varying rates
  - parametric closed form representation of baseline hazard
  - finite no. of parameters
- Makes it really easy to use in calculations of
  - expected residual life time
  - state occupancy probabilities in multistate models
  - ...

Follow-up on several timescales

- The risk-time is the same on all timescales
- Only need the entry point on each time scale:
  - Age at entry.
  - Date of entry.
  - Time since treatment at entry.
    — if time of treatment is the entry, this is 0 for all.
- Response variable in analysis of rates:

$$(d, y) \qquad (\text{event}, \text{duration})$$

- Covariates in analysis of rates:
  - timescales
  - other (fixed) measurements

Follow-up data in `Epi` — `Lexis` objects

A follow-up study:

```
> round( th, 2 )
    id sex birthdat contrast injecdat volume exitdat exitstat
1    1   2  1916.61        1  1938.79     22 1976.79        1
2  640   2  1896.23        1  1945.77     20 1964.37        1
3 3425   1  1886.97        2  1955.18      0 1956.59        1
4 4017   2  1936.81        2  1957.61      0 1992.14        2
...
```

Timescales of interest:

- Age
- Calendar time
- Time since injection

Definition of `Lexis` object

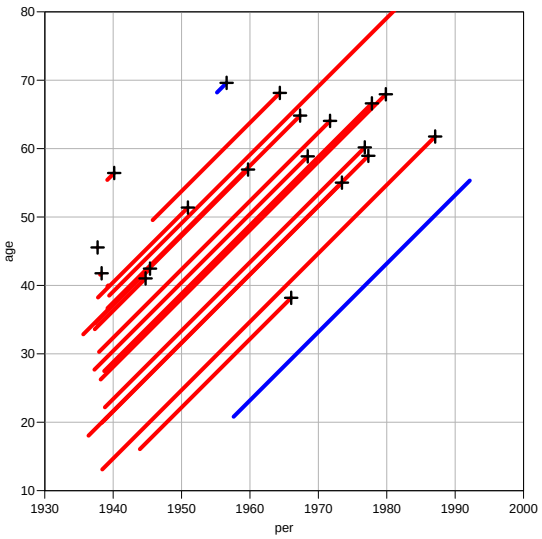
```
> thL <- Lexis( entry = list( age = injecdat-birthdat,
+                             per = injecdat,
+                             tfi = 0 ),
+               exit = list( per = exitdat ),
+         exit.status = as.numeric(exitstat==1),
+               data = th )
```

`entry` is defined on **three** timescales,
but `exit` is only defined on **one** timescale:
Follow-up time is the same on all timescales:

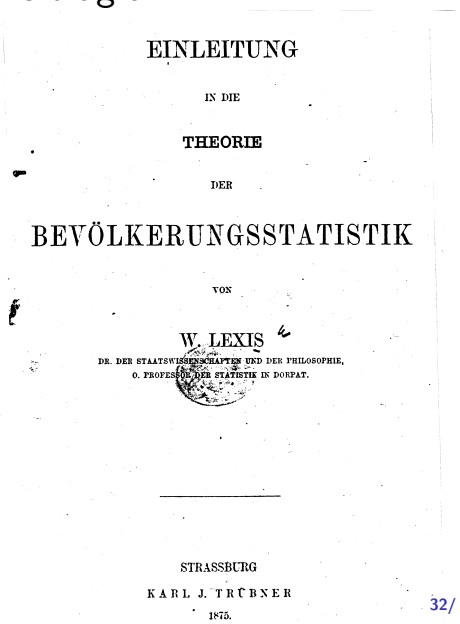
`exitdat - injecdat`

The looks of a `Lexis` object

```
> thL[,1:9]
    age     per tfi lex.dur lex.Cst lex.Xst lex.id
1 22.18 1938.79   0   37.99       0       1      1
2 49.54 1945.77   0   18.59       0       1      2
3 68.20 1955.18   0    1.40       0       1      3
4 20.80 1957.61   0   34.52       0       0      4
...

> summary( thL )
Transitions:
     To
From 0  1 Records:  Events:  Risk time:  Persons:
   0 3 20        23       20      512.59        23
```

Lexis diagram


EINLEITUNG

IN DIE

THEORIE

DER

BEVÖLKERUNGSSTATISTIK

VON

W. LEXIS

DR. DER STAATSWISSENSCHAFTEN UND DER PHILOSOPHIE,
O. PROFESSOR DER STATISTIK IN DORPAT.

STRASSBURG

KARL J. TRÜBNER

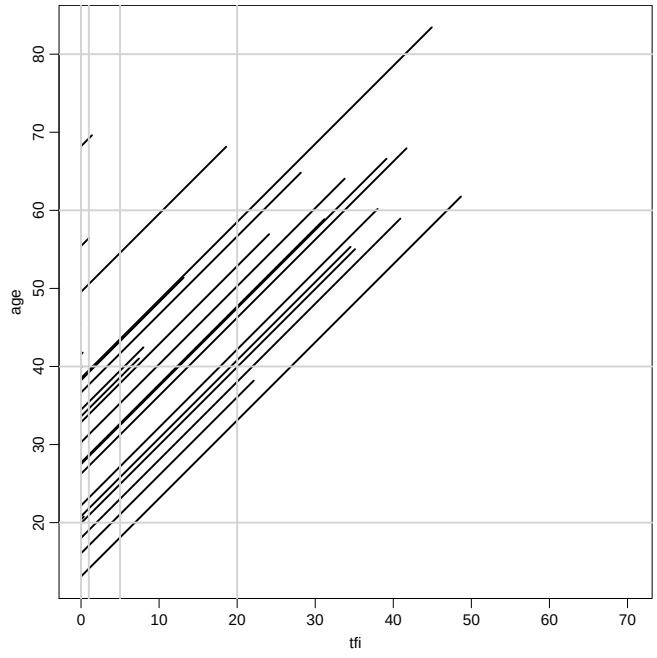
1875.

Splitting follow-up time

```
> spl1 <- splitLexis( thL, breaks=seq(0,100,20),
>                          time.scale="age" )
> round(spl1,1)
   age    per  tfi lex.dur lex.Cst lex.Xst   id sex birthdat contrast injecdat vol
1 22.2 1938.8  0.0    17.8       0       0    1   2    1916.6        1   1938.8
2 40.0 1956.6 17.8    20.0       0       0    1   2    1916.6        1   1938.8
3 60.0 1976.6 37.8     0.2       0       1    1   2    1916.6        1   1938.8
4 49.5 1945.8  0.0    10.5       0       0  640   2    1896.2        1   1945.8
5 60.0 1956.2 10.5     8.1       0       1  640   2    1896.2        1   1945.8
6 68.2 1955.2  0.0     1.4       0       1 3425   1    1887.0        2   1955.2
7 20.8 1957.6  0.0    19.2       0       0 4017   2    1936.8        2   1957.6
8 40.0 1976.8 19.2    15.3       0       0 4017   2    1936.8        2   1957.6
...
```


Split on another timescale

```
> spl2 <- splitLexis( spl1, time.scale="tfi",
                           breaks=c(0,1,5,20,100) )
> round( spl2, 1 )
   lex.id  age    per  tfi lex.dur lex.Cst lex.Xst   id sex birthdat contrast inje
1       1 22.2 1938.8  0.0     1.0       0       0    1   2    1916.6        1    19
2       1 23.2 1939.8  1.0     4.0       0       0    1   2    1916.6        1    19
3       1 27.2 1943.8  5.0    12.8       0       0    1   2    1916.6        1    19
4       1 40.0 1956.6 17.8     2.2       0       0    1   2    1916.6        1    19
5       1 42.2 1958.8 20.0    17.8       0       0    1   2    1916.6        1    19
6       1 60.0 1976.6 37.8     0.2       0       1    1   2    1916.6        1    19
7       2 49.5 1945.8  0.0     1.0       0       0  640   2    1896.2        1    19
8       2 50.5 1946.8  1.0     4.0       0       0  640   2    1896.2        1    19
9       2 54.5 1950.8  5.0     5.5       0       0  640   2    1896.2        1    19
10      2 60.0 1956.2 10.5     8.1       0       1  640   2    1896.2        1    19
11      3 68.2 1955.2  0.0     1.0       0       0 3425   1    1887.0        2    19
12      3 69.2 1956.2  1.0     0.4       0       1 3425   1    1887.0        2    19
13      4 20.8 1957.6  0.0     1.0       0       0 4017   2    1936.8        2    19
14      4 21.8 1958.6  1.0     4.0       0       0 4017   2    1936.8        2    19
15      4 25.8 1962.6  5.0    14.2       0       0 4017   2    1936.8        2    19
16      4 40.0 1976.8 19.2     0.8       0       0 4017   2    1936.8        2    19
17      4 40.8 1977.6 20.0    14.5       0       0 4017   2    1936.8        2    19
```


```
 age  tfi lex.dur lex.Cst lex.Xst
22.2  0.0     1.0       0       0
23.2  1.0     4.0       0       0
27.2  5.0    12.8       0       0
40.0 17.8     2.2       0       0
42.2 20.0    17.8       0       0
60.0 37.8     0.2       0       1
```

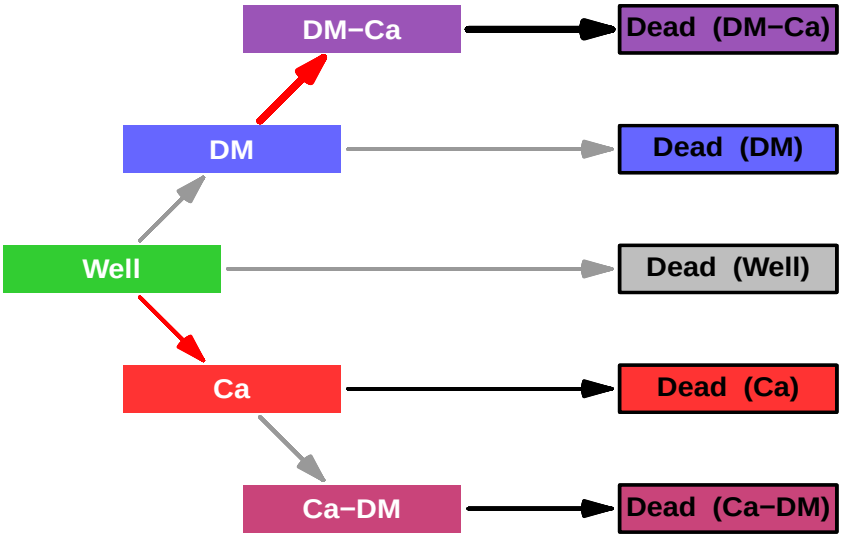
Analysis of results

- d_{it} — events in the variable: `lex.Xst`:
  In the model as response: `lex.Xst==1`
- y_{it} — risk time: `lex.dur` (duration):
  In the model as offset $\log(y)$, `log(lex.dur)`.
- Covariates are:
  - timescales (age, period, time in study)
    — non-linear, continuous effect
  - other variables for this person (constant in each interval).
- If intervals sufficiently small, a very good approximation to a continuously varying rate by using time points from each interval
- And very handy post-processing of results

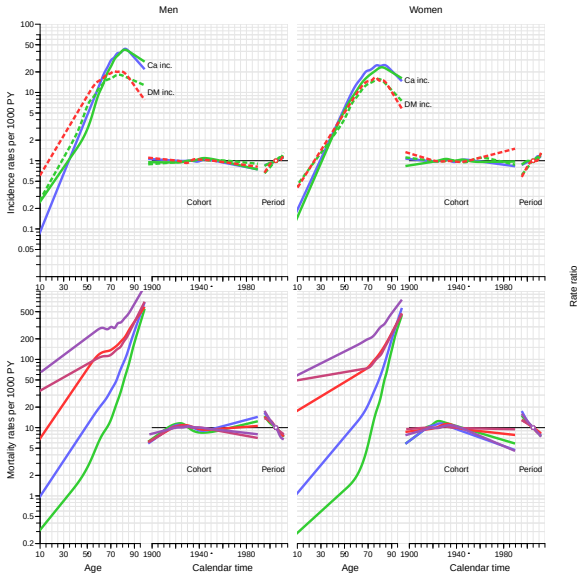
Markov predictions from non-Markov models

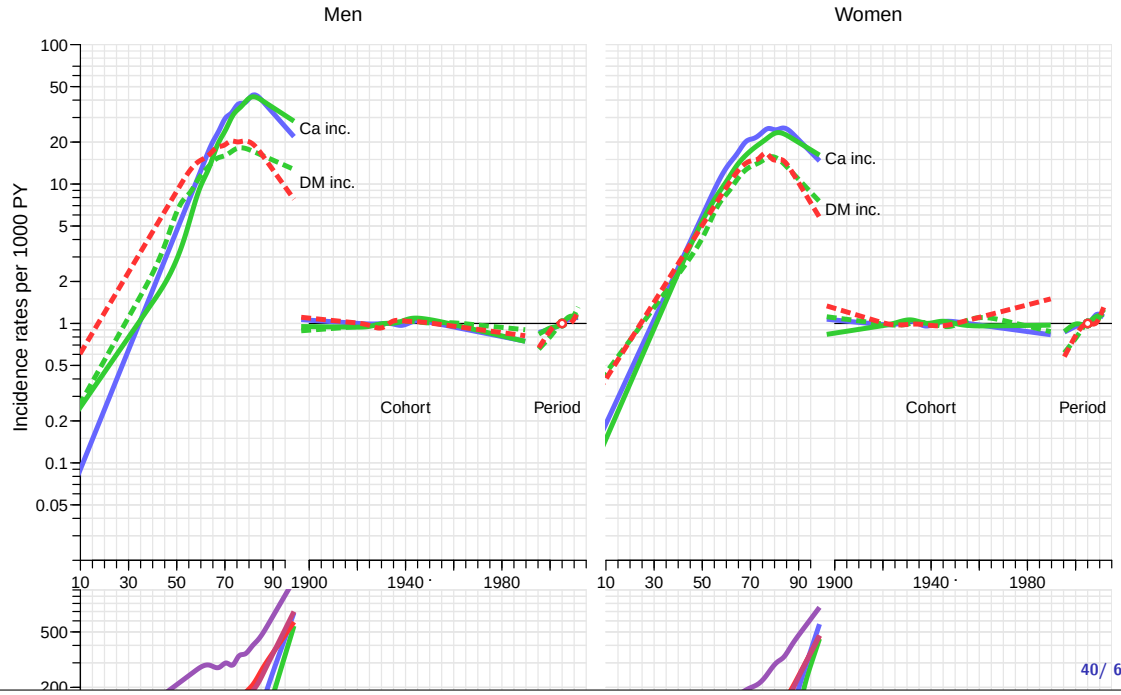
- Model rates in a Lexis diagram (age / calendar time):
  $\lambda(a, t)$
- Aim is summary measures:
  - Expected life time
  - Lifetime probability of disease
  - Lifetime spent diseased
  - ...
- Easy if rates only depend on age
- — so use cross-sectional rates: $\lambda(a, t = T_0)$

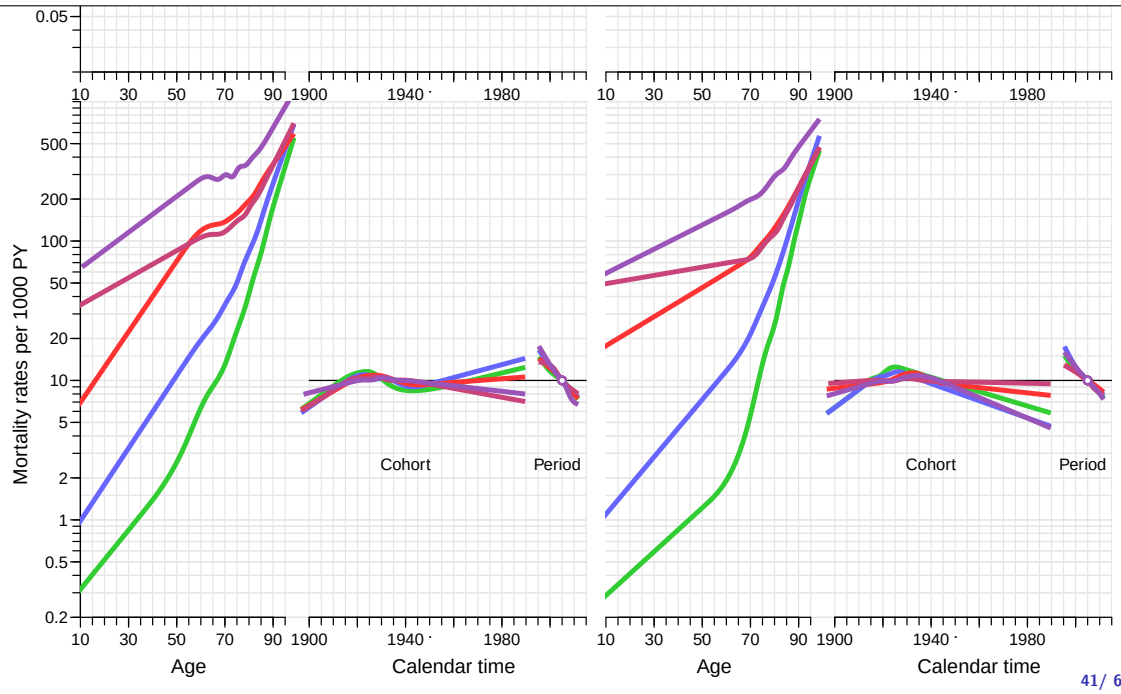
Joint occurrence of Diabetes and Cancer



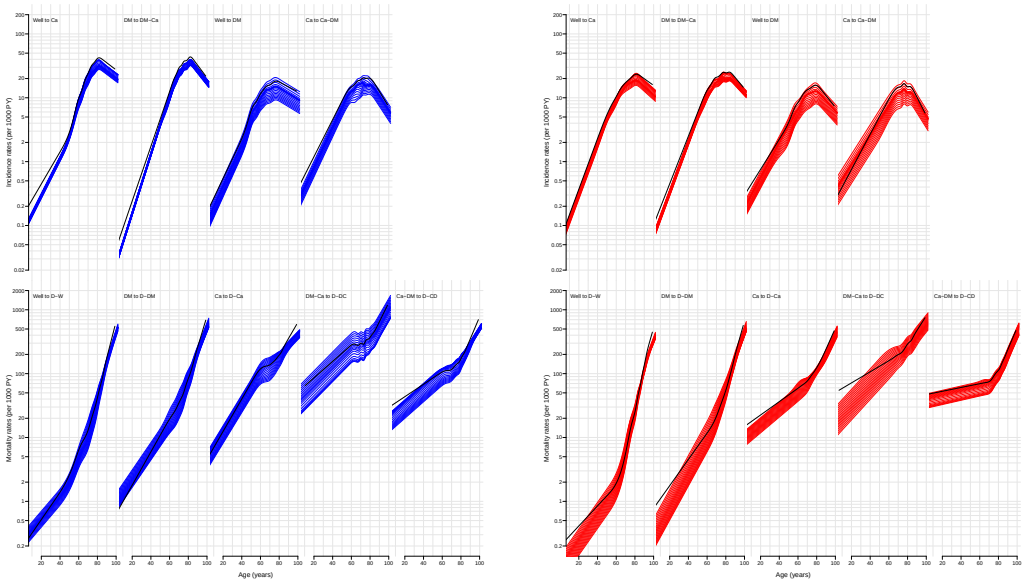
Joint occurrence of Diabetes and Cancer







Predicted rates — cross-sectional 1995–2010



Continuous rates (per 2010)

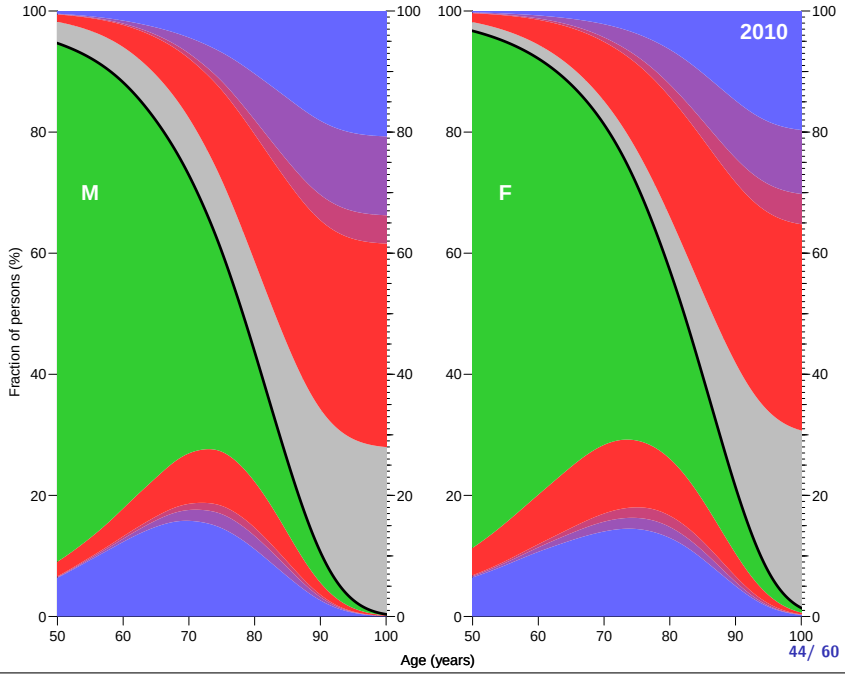
1-month cumulative rates $\rightarrow$ transition probabilities

$$\Big(1-\exp\big(-(\Lambda_1+\Lambda_2+\Lambda_3)\big)\Big)\times\Lambda_i/(\Lambda_1+\Lambda_2+\Lambda_3),\, i=1,2,3$$

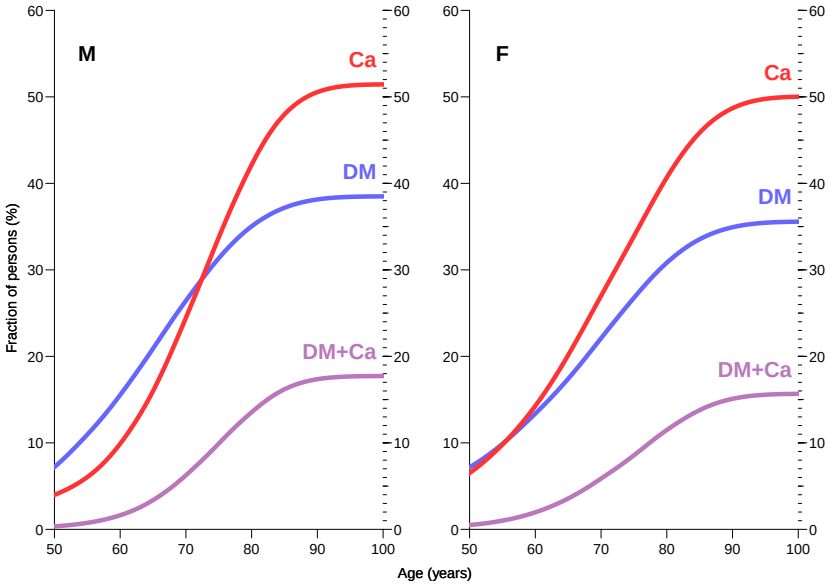
1-month transition probabilities ($\times 10^4$) at age 66 years:

from \ to	Well	DM	DM-Ca	Ca	Ca-DM	D-W	D-DM	D-Ca	D-DC	D-CD	Sum
Well	9966	8	.	13	.	14	.	.	.	.	10000
DM	.	9943	16	.	.	.	41	.	.	.	10000
DM-Ca	.	.	9582	.	.	.	.	.	418	.	10000
Ca	.	.	.	9819	9	.	.	172	.	.	10000
Ca-DM	.	.	.	.	9866	.	.	.	.	134	10000
D-W	.	.	.	.	.	10000	.	.	.	.	10000
D-DM	.	.	.	.	.	.	10000	.	.	.	10000
D-Ca	.	.	.	.	.	.	.	10000	.	.	10000
D-DC	.	.	.	.	.	.	.	.	10000	.	10000
D-CD	.	.	.	.	.	.	.	.	.	10000	10000

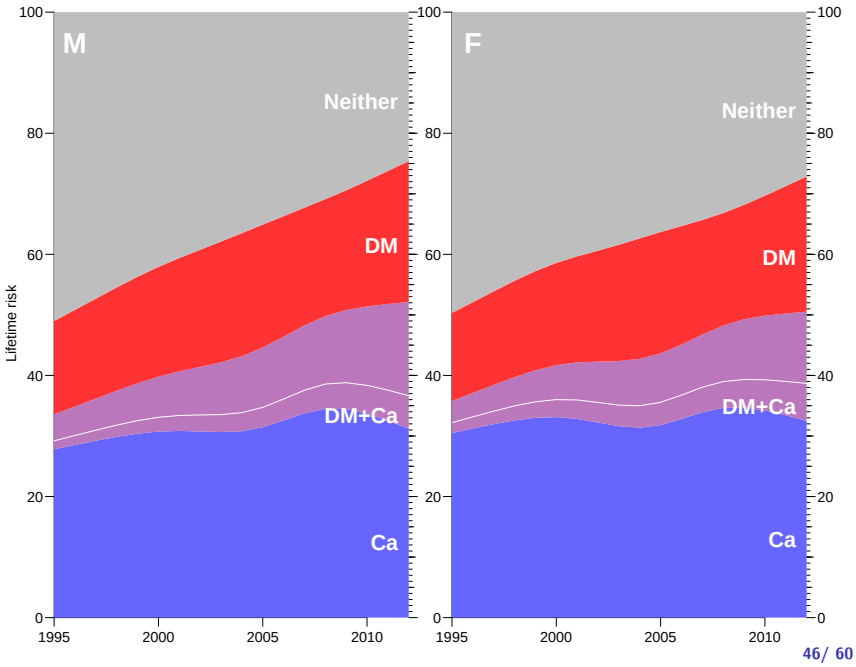
State occupancy probabilities



Lifetime risk



Trend in lifetime risk



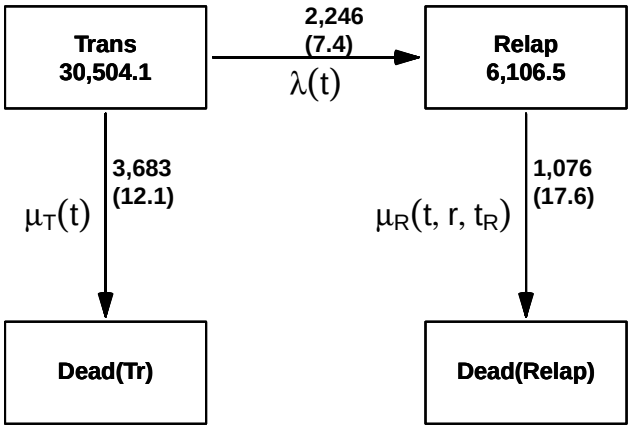
Continuous time rates

- Transition rates between states:
  - based on 1-year tabulation of data
  - age-period-cohort models
  - using smooth effects of age, period and cohort
- Assuming only one transition per interval: small intervals
- State probabilities simple closed-form function of rates
- Numerical integration of closed form functions trivial
- Matrix multiplication trivial

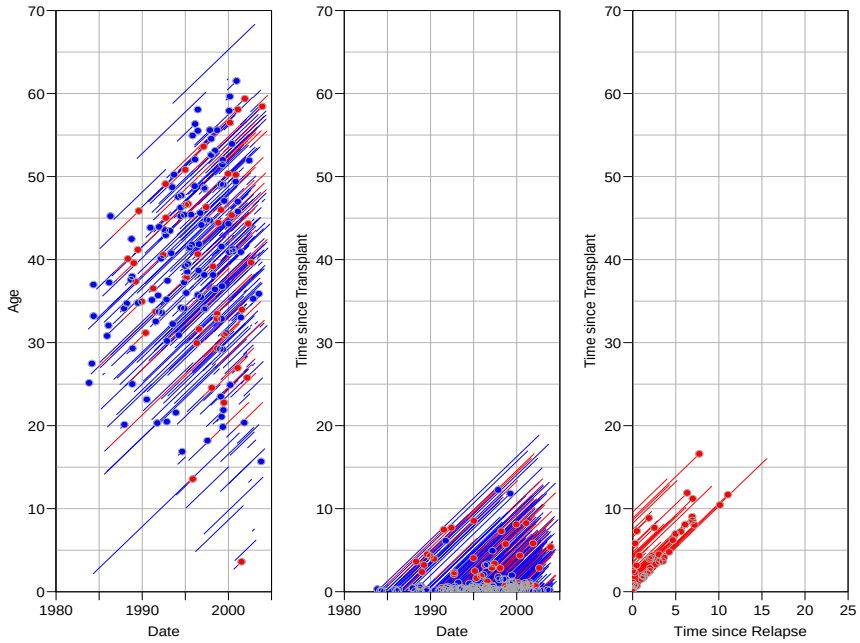
. . . simplified by a parametric form for rates as function of time

EBMT transplant data

Iacobelli & Carstensen: Multistate Models with Multiple Timescales, Stat Med 2013, [3]



other covariates: Age and date at Tx, sex, donor type, CML type



Markov property: Empirical question

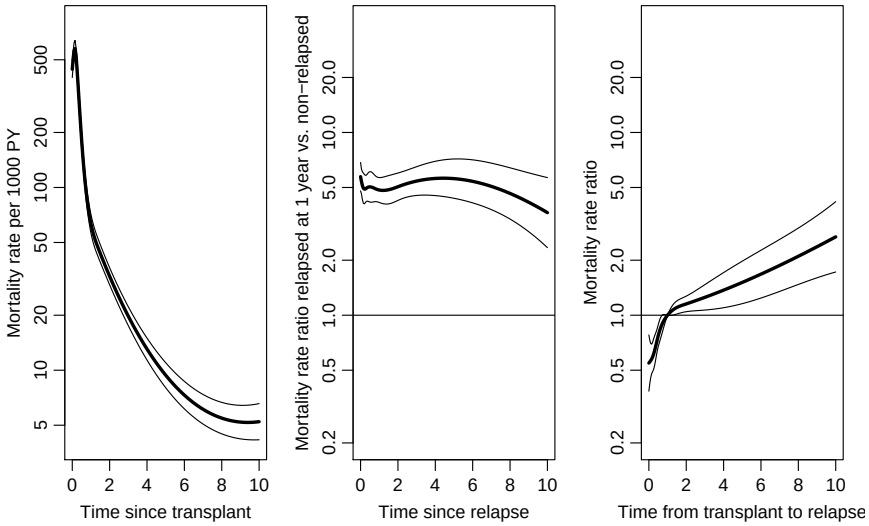
Model for mortality rates:

- t time since transplant
- r time since relapse (if relapsed)
- t_r time from transplant to relapse
- Fit the model for all transitions:
  - split follow-up time
  - fit Poisson model with covariates
  - and spline terms for each time scale.
- `Lexis` machinery from the **Epi** package for **R**
- . . . for representation and manipulation of follow-up data.

Using the Lexis machinery [4, 5]

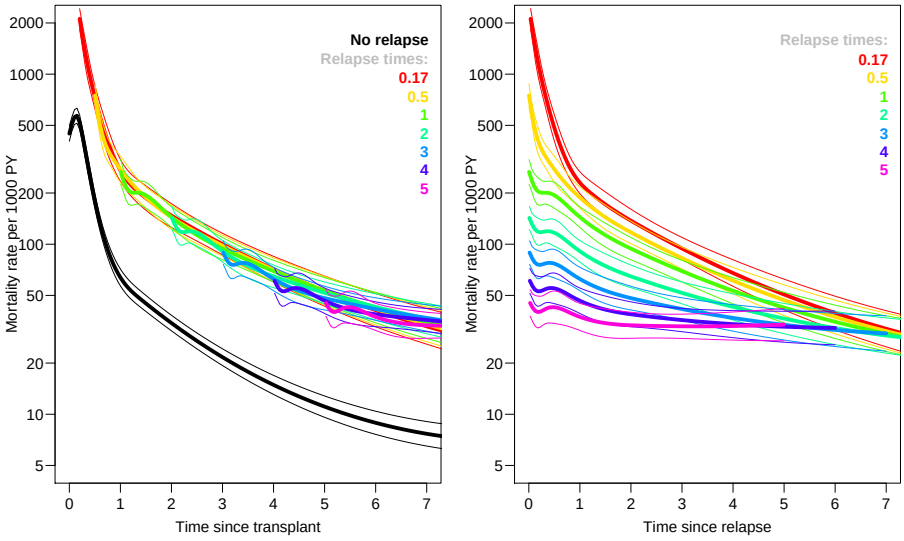
```
cmlT <- Lexis(entry = list(cal = cal.yr(dot),
                           age = cal.yr(dot)-cal.yr(dob),
                           tst = 0),
              exit = list(cal = cal.yr(dof)),
       exit.status = dead,
            states = c("Transplant","Dead"),
              data = cml )

cmlL <- cutLexis( cmlT, cut = cal.yr(cmlT$dor),
                  new.state = "Relapse",
                  new.scale = "tsr",
             precursor.states = "Transplant")

> subset( cmlL, lex.id==151 )[,1:8]

 id     cal   age  tst  tsr lex.dur lex.Cst lex.Xst  covariates
151 1987.28 36.22 0.00  NA    1.87   Trans   Relap  ...
151 1989.16 38.10 1.87   0    4.93   Relap    Dead  ...
```


$$\log(\mu) = h(t) + k(r) + g(t-r) + X\beta$$



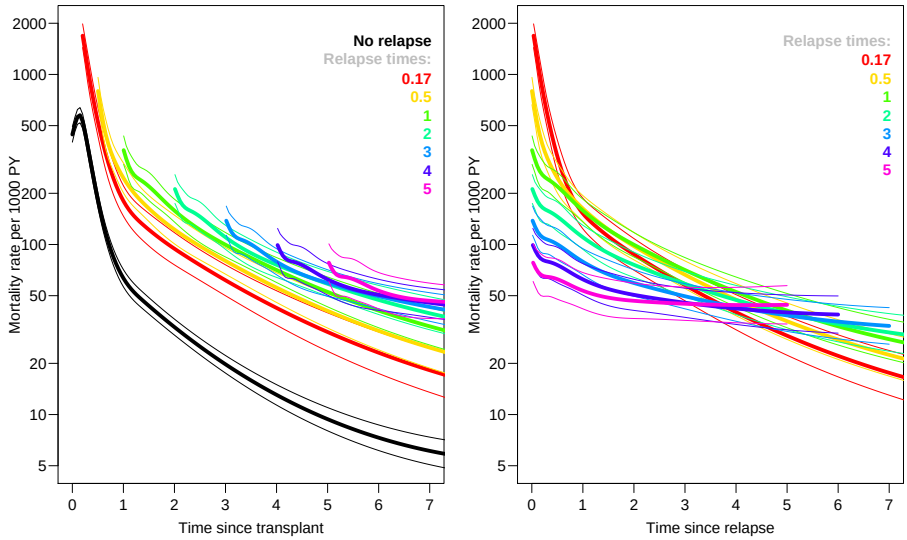
t: time since transplant r: time since relapse

$$\log(\mu) = h(t) + k(r) \qquad + X\beta$$



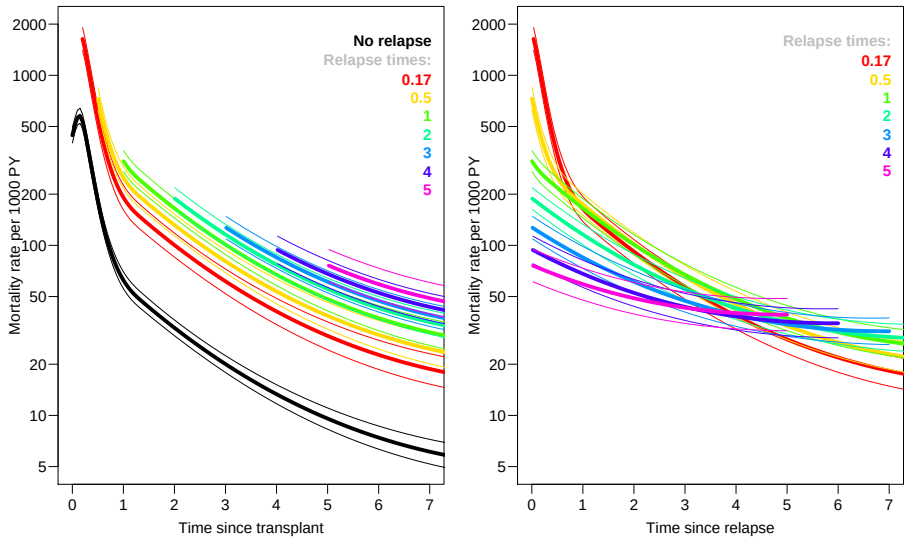
t: time since transplant r: time since relapse

$$\log(\mu) = h(t) + k(r) + g(t - r) + X\beta$$



t: time since transplant r: time since relapse

$$\log(\mu) = h(t) \qquad + g(t - r) + X\beta$$



t: time since transplant r: time since relapse

Model summary

- Mortality of relapsed patients depends on **when** they relapsed.
- We also checked if the mortality depended on **time since** they relapsed.
  It did not.
- **Note:** It is an **empirical** question what timescales to use.
- **Note:** In order to compute probabilities, we need a model for the relapse rates (λ) in addition to the mortality rates (μ_T, μ_R)
- . . . unfortunately not a Markov model

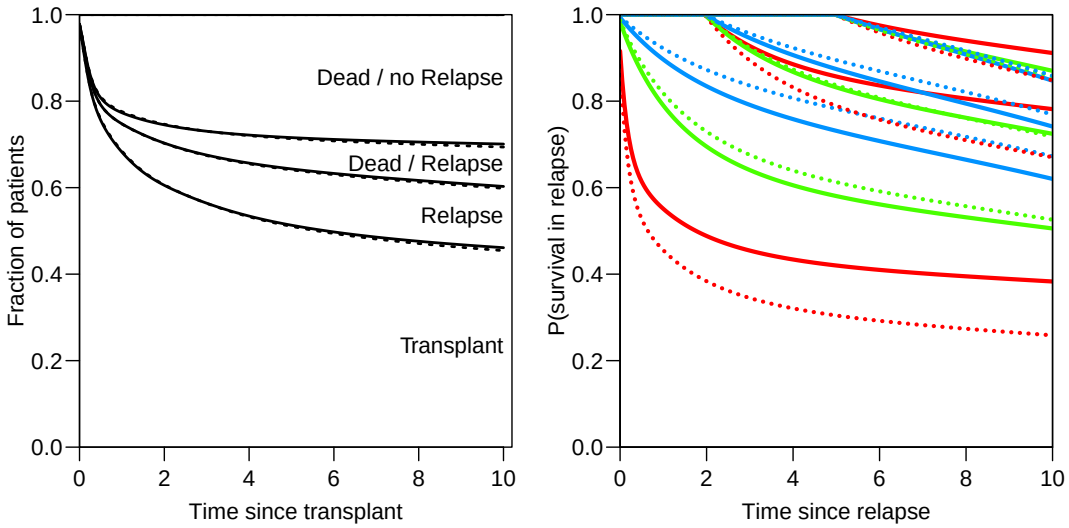
Not Markov: the hard way

$$\mathrm{P}\{\text{T at } t\} = \exp\Big(-\int_0^t \lambda(s) + \mu_T(s)\,\mathrm{d}s\Big)$$

$$\mathrm{P}\{\text{D(T) at } t\} = \int_0^t \mu_T(s)\exp\Big(-\int_0^s \lambda(u) + \mu_T(u)\,\mathrm{d}u\Big)\,\mathrm{d}s$$

$$\begin{aligned}\mathrm{P}\{\text{R at } t\} &= \int_0^t \mathrm{P}\{\text{Relapsed at } s\} \\ &\quad \times \mathrm{P}\{\text{Survive in Relapse from } s \text{ to } t\}\,\mathrm{d}s \\ &= \int_0^t \lambda(s)\exp\Big(-\int_0^s \lambda(u) + \mu_T(u)\,\mathrm{d}u\Big) \\ &\quad \times \exp\Big(-\int_s^t \mu_R(u,s)\,\mathrm{d}u\Big)\,\mathrm{d}s\end{aligned}$$

$$\mathrm{P}\{\text{D(R) at } t\} = 1 - \mathrm{P}\{\text{T at } t\} - \mathrm{P}\{\text{D(T) at } t\} - \mathrm{P}\{\text{R at } t\}$$



Dotted lines: Markov model, time since transplant
Full lines: + time from Tx to Rel for the μ_R

Rel at: 2 mth, 1 y, 3 y

Summary & Conclusions

- The world is continuous
- Time effect likely to be smooth
- A single time scale is rarely sufficient
- Different timescales require joint reporting
- Continuous time formulae easiest to handle:
  - Parametric form of time-effects allow direct implementation of probability theory
  - Choice of time scales is an **empirical** problem
- Non/Semi-parametric survival model not well suited for this
- Stick to this world: Fewer tables — more graphs!

Thanks for your attention

References


P. K. Andersen and N. Keiding.
Interpretability and importance of functionals in competing risks and multistate models.
Stat Med, 31:1074–1088, 2012.


S. Suissa.
Immortal time bias in pharmaco-epidemiology.
Am. J. Epidemiol., 167:492–499, Feb 2008.


S. Iacobelli and B. Carstensen.
Multiple time scales in multi-state models.
Stat Med, 32(30):5315–5327, Dec 2013.


Martyn Plummer and Bendix Carstensen.
Lexis: An R class for epidemiological studies with long-term follow-up.
Journal of Statistical Software, 38(5):1–12, 1 2011.


Bendix Carstensen and Martyn Plummer.
Using Lexis objects for multi-state models in R.
Journal of Statistical Software, 38(6):1–18, 1 2011.