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

Knibbs lecture, SSAI, Canberra, Australia, 24 November 2015

Inference in Multistate models

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

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

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

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

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].

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].
- Wrongly including persons' follow-up in the wrong state (namely the one reached some time in the future).

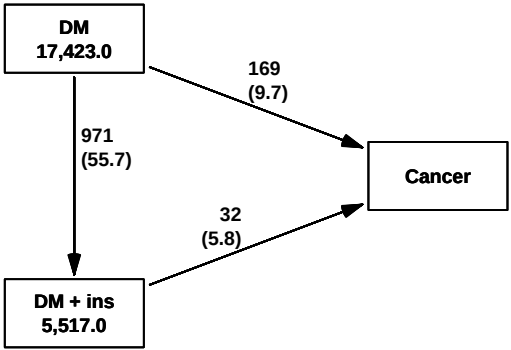
Conditioning on the future

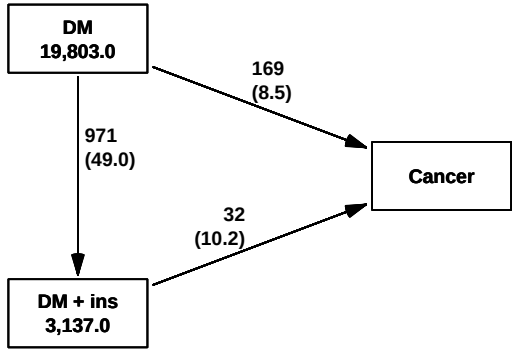
- . . . also known as "Immortal time bias", see *e.g.* S. Suissa: Immortal time bias in pharmaco-epidemiology, *Am. J. Epidemiol*, 2008 [2].
- Wrongly including persons' follow-up in the wrong state (namely the one reached some time in the future).
- Frequently caused by classification of **persons**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

Immortal time bias

Yang *et al.*:
Associations of hyperglycemia and insulin usage with the risk of cancer in type 2 diabetes: the Hong Kong diabetes registry, *Diabetes*, 2010 [3]
. . . found that the RR of cancer associated with insulin use among diabetes patients were 0.22 — very small indeed.

This was challenged [4] because person-years enumeration was possible from the published tables.





$5.8/9.7 = 0.60$

$10.2/8.5 = 1.20$

Why these mistakes?

- Time is absent from survival analysis results

Why these mistakes?

- Time is absent from survival analysis results
- Time is taken to be a **response** variable observed for each person

Why these mistakes?

- Time is absent from survival analysis results
- Time is taken to be a **response** variable observed for each person
- Unit of analysis seems to be the person

Why these mistakes?

- Time is absent from survival analysis results
- Time is taken to be a **response** variable observed for each person
- Unit of analysis seems to be the person
- **Persons** classified by exposure

Why these mistakes?

- Time is absent from survival analysis results
- Time is taken to be a **response** variable observed for each person
- Unit of analysis seems to be the person
- **Persons** classified by exposure
- The **real** unit of observation should be person-time

Time

- Time is a **covariate** — determinant of rates

Time

- Time is a **covariate** — determinant of rates
- **Response** variable in survival / follow-up is bivariate:

Time

- Time is a **covariate** — determinant of rates
- **Response** variable in survival / follow-up is bivariate:
  - **Differences** on the timescale (**risk** time, "exposure")

Time

- Time is a **covariate** — determinant of rates
- **Response** variable in survival / follow-up is bivariate:
  - **Differences** on the timescale (**risk** time, "exposure")
  - **Events**

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:

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"

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.

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

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

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 and disease duration.

“Stick to this world”

In the paper by Andersen & Keiding this is primarily aimed at the use of “net survival”,

“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-exhustive exit rate from a state.

"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-exhustive exit rate from a state.

Corresponds to the survival probability in the situation where:

1. all other causes of death are absent

"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-exhustive exit rate from a state.

Corresponds to the survival probability in the situation where:

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

"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-exhustive exit rate from a state.

Corresponds to the survival probability in the situation where:

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

"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-exhustive exit rate from a state.

Corresponds to the 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

- Do not make predictions based on unrealistic assumptions:

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or
  1. Smallpox is eradicated

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or
  1. Smallpox is eradicated
  2. . . . yet mortality remains the same

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or
  1. Smallpox is eradicated
  2. . . . yet mortality remains the same
- I postulate a specific feature of "this world":

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or
  1. Smallpox is eradicated
  2. . . . yet mortality remains the same
- I postulate a specific feature of "this world":
- — it is **continuous**

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or
  1. Smallpox is eradicated
  2. . . . yet mortality remains the same
- I postulate a specific feature of "this world":
- — it is **continuous**
- — in particular, death and disease rates vary **smoothly** by

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or
  1. Smallpox is eradicated
  2. . . . yet mortality remains the same
- I postulate a specific feature of "this world":
- — it is **continuous**
- — in particular, death and disease rates vary **smoothly** by
  - age

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or
  1. Smallpox is eradicated
  2. . . . yet mortality remains the same
- I postulate a specific feature of "this world":
- — it is **continuous**
- — in particular, death and disease rates vary **smoothly** by
  - age
  - calendar time

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or
  1. Smallpox is eradicated
  2. . . . yet mortality remains the same
- I postulate a specific feature of "this world":
- — it is **continuous**
- — in particular, death and disease rates vary **smoothly** by
  - age
  - calendar time
  - disease duration

Sticking to this world

- Do not make predictions based on unrealistic assumptions:
  1. Mortality is 0
  2. Cancer rates as now
- or
  1. Smallpox is eradicated
  2. . . . yet mortality remains the same
- I postulate a specific feature of "this world":
- — it is **continuous**
- — in particular, 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.

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.

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)

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 it is less clear
  — t is but a covariate that varies within individual.

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 it is less clear
  — t is but a covariate that varies within individual.
- Cox's approach profiles $\lambda_0(t)$ out.

The Cox-likelihood as profile likelihood

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

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

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:

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

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

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

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — much larger than the original dataset

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — much larger than the original dataset
- — each individual contributes many empirical rates

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — much larger than the original dataset
- — each individual contributes many empirical rates
- (one per risk-set contribution in Cox-modelling)

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — much larger than the original dataset
- — each individual contributes many empirical rates
- (one per risk-set contribution in Cox-modelling)
- From each empirical rate we get:

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — much larger than the original dataset
- — each individual contributes many empirical rates
- (one per risk-set contribution in Cox-modelling)
- From each empirical rate we get:
  - Poisson-response d

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — much larger than the original dataset
- — 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

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — much larger than the original dataset
- — 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
  - Covariate value for the timescale
    (time since entry, current age, current date, ...)

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — much larger than the original dataset
- — 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
  - Covariate value for the timescale (time since entry, current age, current date, ...)
  - other covariates

Splitting the dataset

- The Poisson approach needs a dataset of empirical rates (d, y) with suitably small values of y.
- — much larger than the original dataset
- — 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
  - Covariate value for the timescale (time since entry, current age, current date, ...)
  - other covariates
- Modelling is by standard `glm` Poisson

Example: Mayo Clinic lung cancer

- Survival after lung cancer

Example: Mayo Clinic lung cancer

- Survival after lung cancer
- Covariates:

Example: Mayo Clinic lung cancer

- Survival after lung cancer
- Covariates:
  - Age at diagnosis

Example: Mayo Clinic lung cancer

- Survival after lung cancer
- Covariates:
  - Age at diagnosis
  - Sex

Example: Mayo Clinic lung cancer

- Survival after lung cancer
- Covariates:
  - Age at diagnosis
  - Sex
  - Time since diagnosis

Example: Mayo Clinic lung cancer

- Survival after lung cancer
- Covariates:
  - Age at diagnosis
  - Sex
  - Time since diagnosis
- Cox model

Example: Mayo Clinic lung cancer

- Survival after lung cancer
- Covariates:
  - Age at diagnosis
  - Sex
  - Time since diagnosis
- Cox model
- Split data:

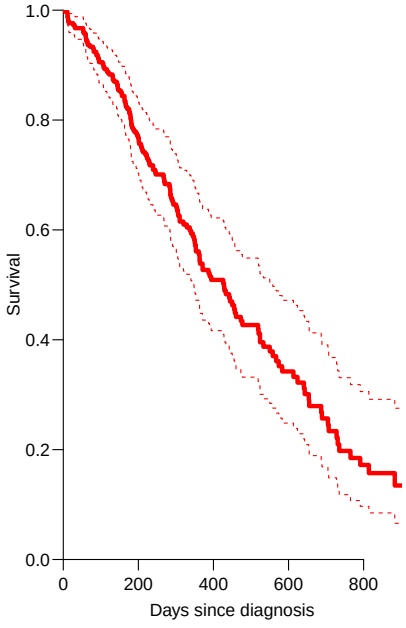
Example: Mayo Clinic lung cancer

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

Example: Mayo Clinic lung cancer

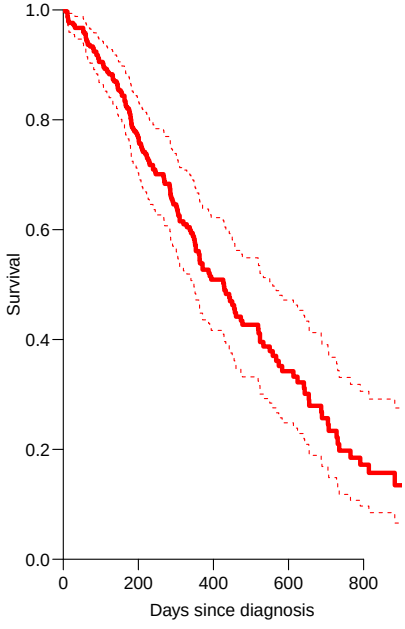
- Survival after lung cancer
- Covariates:
  - Age at diagnosis
  - Sex
  - Time since diagnosis
- Cox model
- Split data:
  - Poisson model, time as factor
  - Poisson model, time as spline

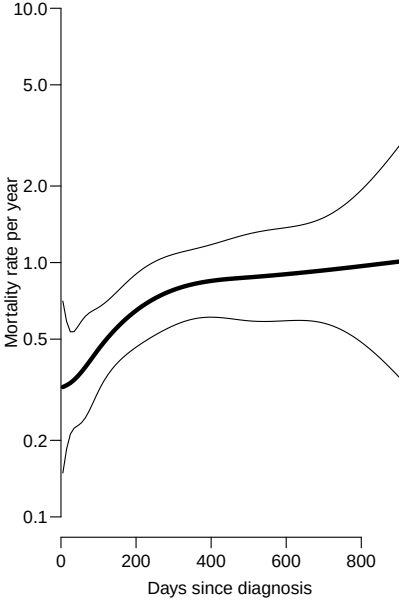
Mayo Clinic lung cancer 60 year old woman

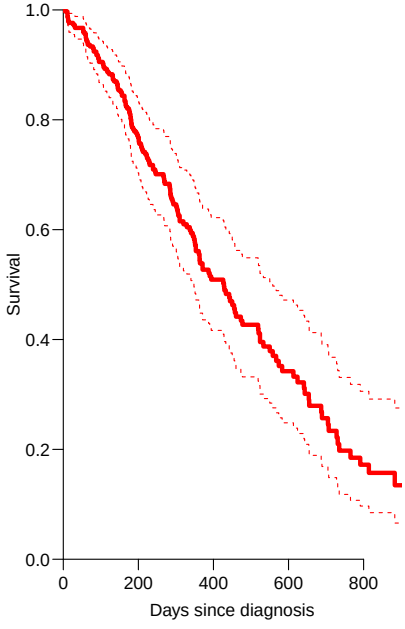


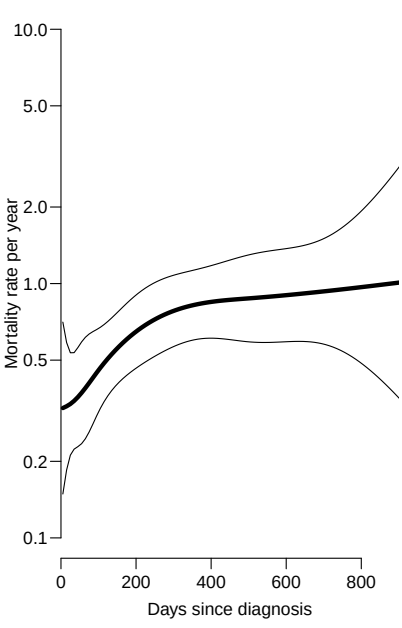
Example: Mayo Clinic lung cancer I

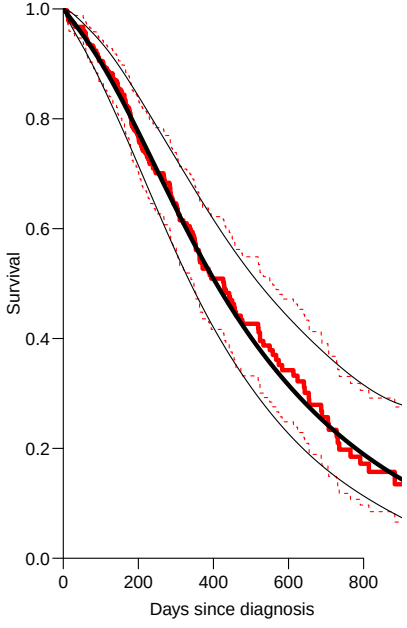
```
> round( cmp, 5 )
                    age    2.5%   97.5%     sex    2.5%   97.5%
Cox             1.01716 0.99894 1.03571 0.59896 0.43137 0.83165
Poisson-factor  1.01716 0.99894 1.03571 0.59896 0.43137 0.83165
Poisson-spline  1.01619 0.99803 1.03468 0.59983 0.43199 0.83287
```

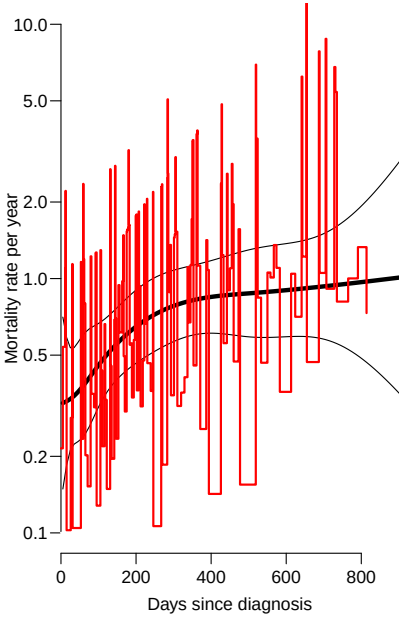

1.0
0.8
0.6
0.4
0.2
0.0
Survival
0
200
400
600
800
Days since diagnosis

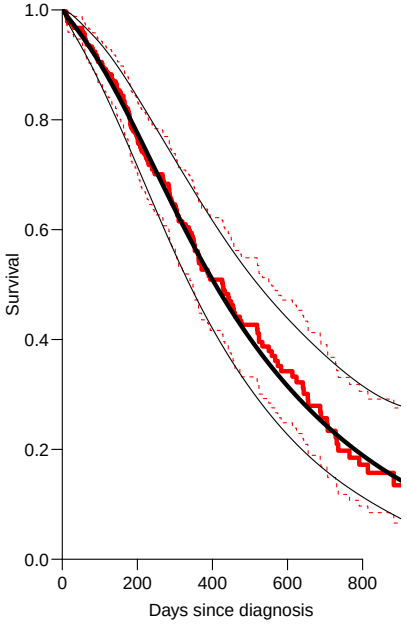

10.0
5.0
2.0
1.0
0.5
0.2
0.1
Mortality rate per year
0
200
400
600
800
Days since diagnosis

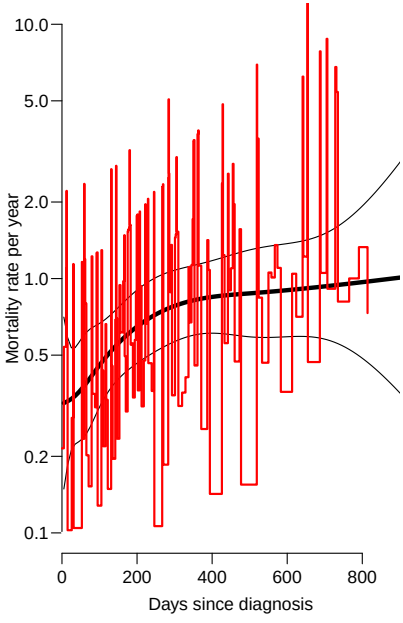

1.0
0.8
0.6
0.4
0.2
0.0
Survival
0
200
400
600
800
Days since diagnosis

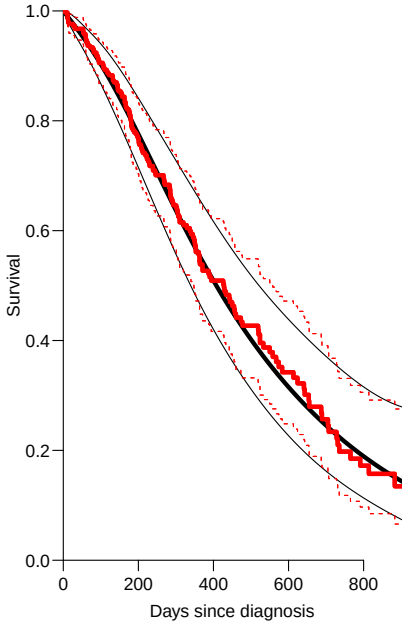

10.0
5.0
2.0
1.0
0.5
0.2
0.1
Mortality rate per year
0
200
400
600
800
Days since diagnosis


1.0
0.8
0.6
0.4
0.2
0.0
Survival
0
200
400
600
800
Days since diagnosis


10.0
5.0
2.0
1.0
0.5
0.2
0.1
Mortality rate per year
0
200
400
600
800
Days since diagnosis


1.0
0.8
0.6
0.4
0.2
0.0
Survival
0
200
400
600
800
Days since diagnosis


10.0
5.0
2.0
1.0
0.5
0.2
0.1
Mortality rate per year
0
200
400
600
800
Days since diagnosis


1.0
0.8
0.6
0.4
0.2
0.0
Survival
0
200
400
600
800
Days since diagnosis

```
> 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))
```

What the Cox-model really is

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

- dividing time very finely and

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.

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.
- $\Rightarrow$ difficult to access the baseline hazard.

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.
- $\Rightarrow$ difficult to access the baseline hazard.
- $\Rightarrow$ uninitiated tempted to show survival curves where irrelevant

Modeling in this world

- Replace the α_ts by a parametric function $f(t)$ with a limited number of parameters, for example:

Modeling in this world

- Replace the α_ts by a parametric function $f(t)$ with a limited number of parameters, for example:
  - Piecewise constant

Modeling in 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)

Modeling in 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

Modeling in 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":

Modeling in 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

Modeling in 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

Modeling in 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

Modeling in 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

Modeling in 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

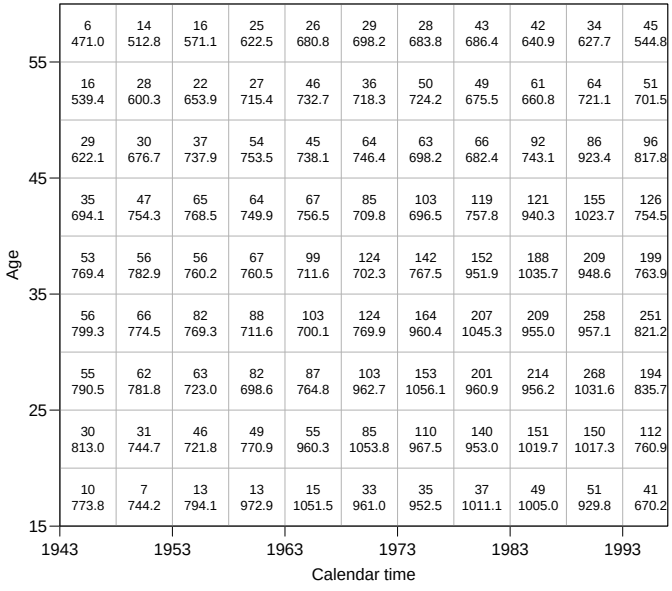
Modeling in this world

- 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

Modeling in 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
  - ...

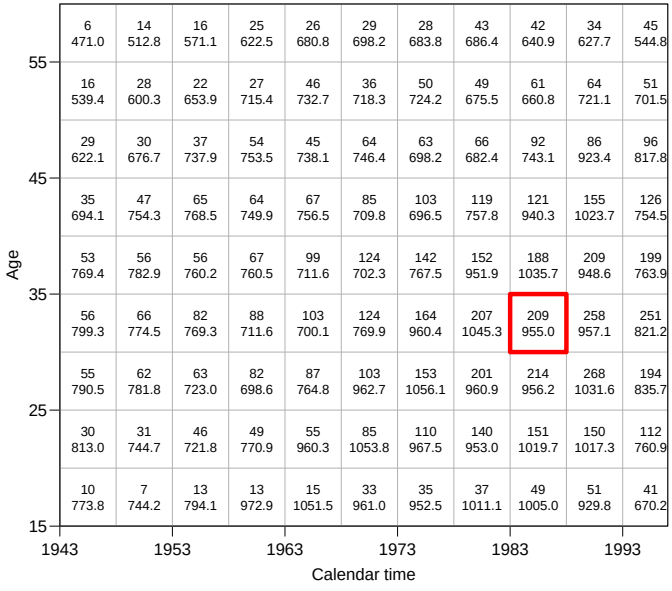
Tabulation of register data



Testis cancer cases in Denmark.

Male person-years in Denmark.

Tabulation of register data

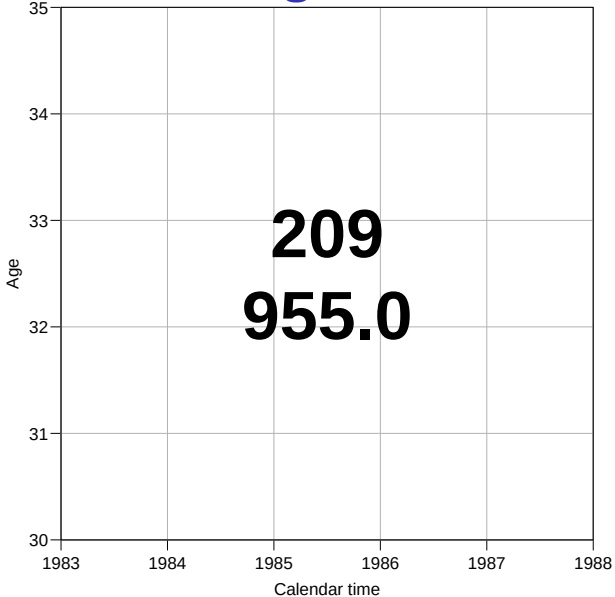


Age	1943	1953	1963	1973	1983	1993
55	6 471.0	14 512.8	16 571.1	25 622.5	26 680.8	29 698.2	28 683.8	43 686.4	42 640.9	34 627.7	45 544.8
45	16 539.4	28 600.3	22 653.9	27 715.4	46 732.7	36 718.3	50 724.2	49 675.5	61 660.8	64 721.1	51 701.5
35	29 622.1	30 676.7	37 737.9	54 753.5	45 738.1	64 746.4	63 698.2	66 682.4	92 743.1	86 923.4	96 817.8
25	35 694.1	47 754.3	65 768.5	64 749.9	67 756.5	85 709.8	103 696.5	119 757.8	121 940.3	155 1023.7	126 754.5
15	53 769.4	56 782.9	56 760.2	67 760.5	99 711.6	124 702.3	142 767.5	152 951.9	188 1035.7	209 948.6	199 763.9
	56 799.3	66 774.5	82 769.3	88 711.6	103 700.1	124 769.9	164 960.4	207 1045.3	209 955.0	258 957.1	251 821.2
	55 790.5	62 781.8	63 723.0	82 698.6	87 764.8	103 962.7	153 1056.1	201 960.9	214 956.2	268 1031.6	194 835.7
	30 813.0	31 744.7	46 721.8	49 770.9	55 960.3	85 1053.8	110 967.5	140 953.0	151 1019.7	150 1017.3	112 760.9
	10 773.8	7 744.2	13 794.1	13 972.9	15 1051.5	33 961.0	35 952.5	37 1011.1	49 1005.0	51 929.8	41 670.2

Testis cancer cases in Denmark.

Male person-years in Denmark.

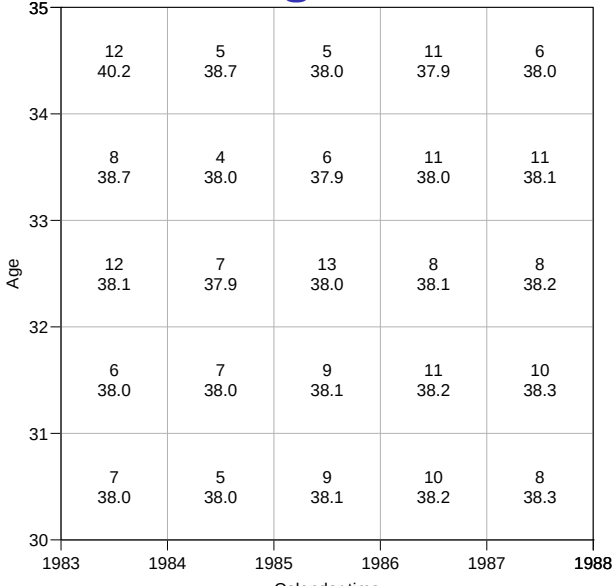
Tabulation of register data



Testis cancer cases in Denmark.

Male person-years in Denmark.

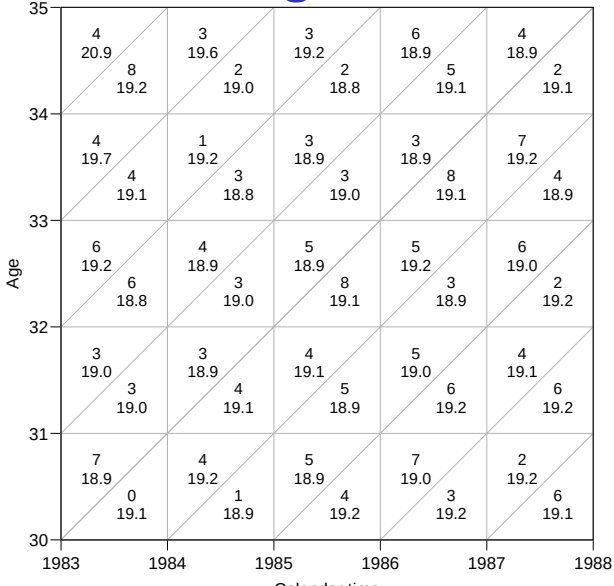
Tabulation of register data



Testis cancer cases in Denmark.

Male person-years in Denmark.

Tabulation of register data

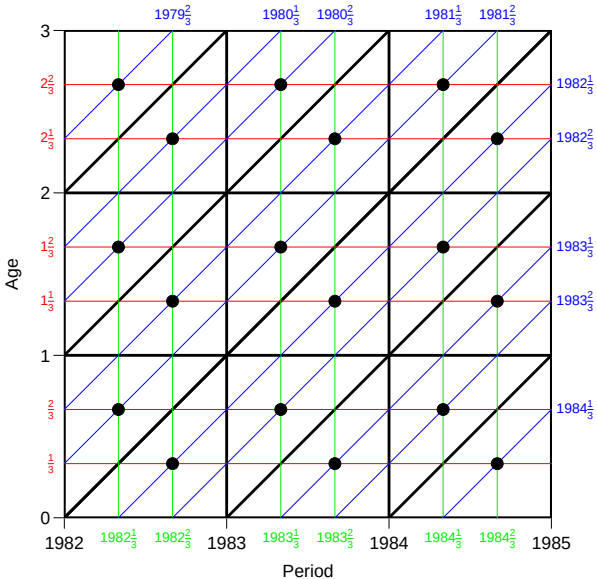


Testis cancer cases in Denmark.

Male person-years in Denmark.

Subdivision by year of birth (cohort).

Tabulation by age, period and cohort



Gives triangular sets with differing mean age, period and cohort:

These correct midpoints for age, period and cohort must be used in modelling.

Model for triangular data

- One parameter per distinct value on each timescale.

Model for triangular data

- One parameter per distinct value on each timescale.
- Example: 4 age-classes and 4 periods would give 32 observations and 30 parameters

Model for triangular data

- One parameter per distinct value on each timescale.
- Example: 4 age-classes and 4 periods would give 32 observations and 30 parameters
  - 8 age parameters

Model for triangular data

- One parameter per distinct value on each timescale.
- Example: 4 age-classes and 4 periods would give 32 observations and 30 parameters
  - 8 age parameters
  - 8 period parameters

Model for triangular data

- One parameter per distinct value on each timescale.
- Example: 4 age-classes and 4 periods would give 32 observations and 30 parameters
  - 8 age parameters
  - 8 period parameters
  - 14 cohort parameters

Model for triangular data

- One parameter per distinct value on each timescale.
- Example: 4 age-classes and 4 periods would give 32 observations and 30 parameters
  - 8 age parameters
  - 8 period parameters
  - 14 cohort parameters
- Model:

$$\log(\lambda_{ap}) = \alpha_a + \beta_p + \gamma_c$$

Model for triangular data

- One parameter per distinct value on each timescale.
- Example: 4 age-classes and 4 periods would give 32 observations and 30 parameters
  - 8 age parameters
  - 8 period parameters
  - 14 cohort parameters
- Model:

$$\log(\lambda_{ap}) = \alpha_a + \beta_p + \gamma_c$$

- . . . only 26 parameters identifiable.

Problem: Disconnected design!

Log-likelihood contribution from one triangle:

$$D_{ap}\log(\lambda_{ap}) - \lambda_{ap}\, Y_{ap} = D_{ap}(\alpha_a + \beta_p + \gamma_c) - \exp(\alpha_a + \beta_p + \gamma_c)\, Y_{ap}$$

Problem: Disconnected design!

Log-likelihood contribution from one triangle:

$$D_{ap}\log(\lambda_{ap}) - \lambda_{ap}\, Y_{ap} = D_{ap}(\alpha_a + \beta_p + \gamma_c) - \exp(\alpha_a + \beta_p + \gamma_c)\, Y_{ap}$$

The total log-likelihood can be separated:

$$\sum_{a,p \in \triangledown} D_{ap}\log(\lambda_{ap}) - \lambda_{ap}\, Y_{ap} + \sum_{a,p \in \triangle} D_{ap}\log(\lambda_{ap}) - \lambda_{ap}\, Y_{ap}$$

Problem: Disconnected design!

Log-likelihood contribution from one triangle:

$$D_{ap}\log(\lambda_{ap}) - \lambda_{ap}\, Y_{ap} = D_{ap}(\alpha_a + \beta_p + \gamma_c) - \exp(\alpha_a + \beta_p + \gamma_c)\, Y_{ap}$$

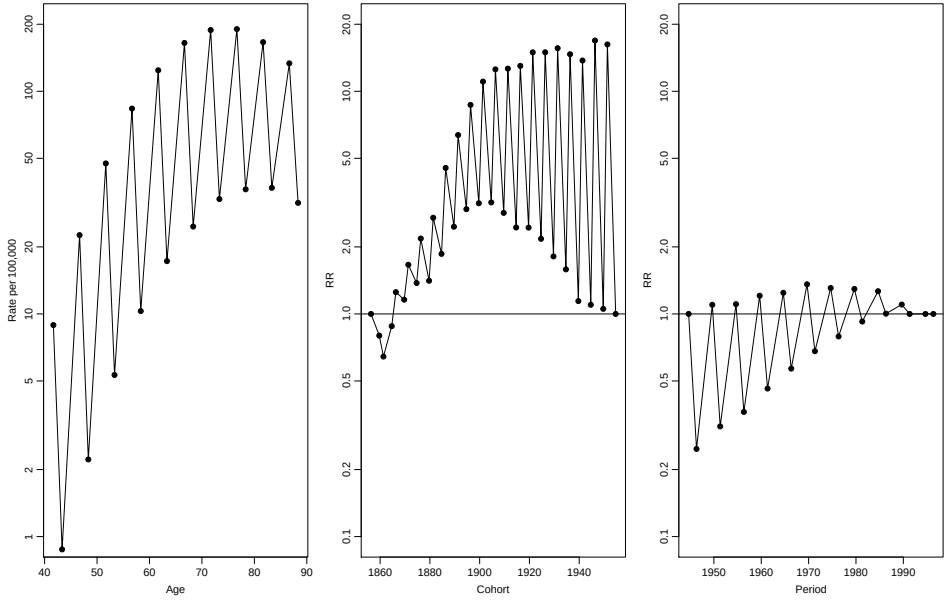
The total log-likelihood can be separated:

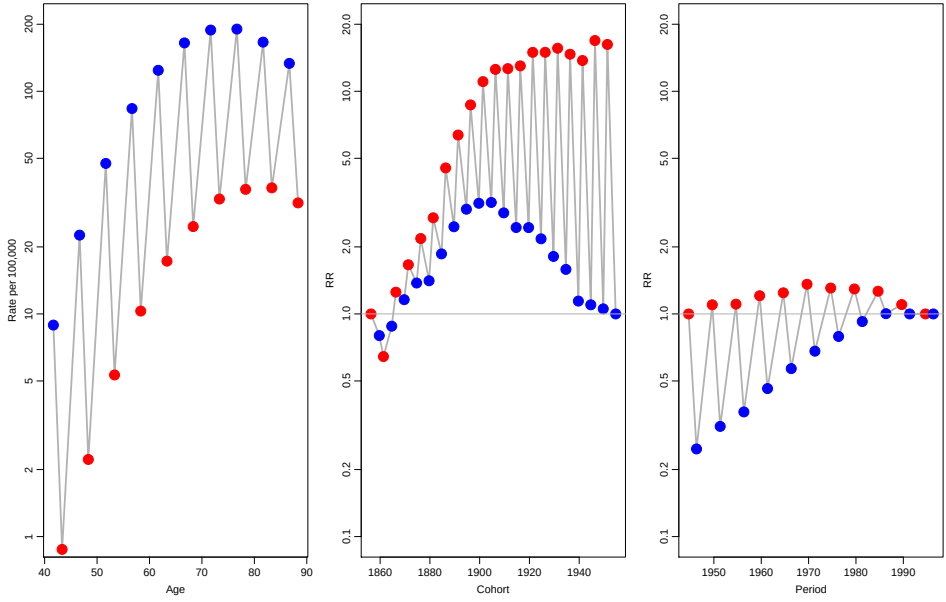
$$\sum_{a,p \in \triangledown} D_{ap}\log(\lambda_{ap}) - \lambda_{ap}\, Y_{ap} + \sum_{a,p \in \triangle} D_{ap}\log(\lambda_{ap}) - \lambda_{ap}\, Y_{ap}$$

No common parameters between terms — two separate models:
One for upper triangles, one for lower.

Illustration by Danish lung cancer data

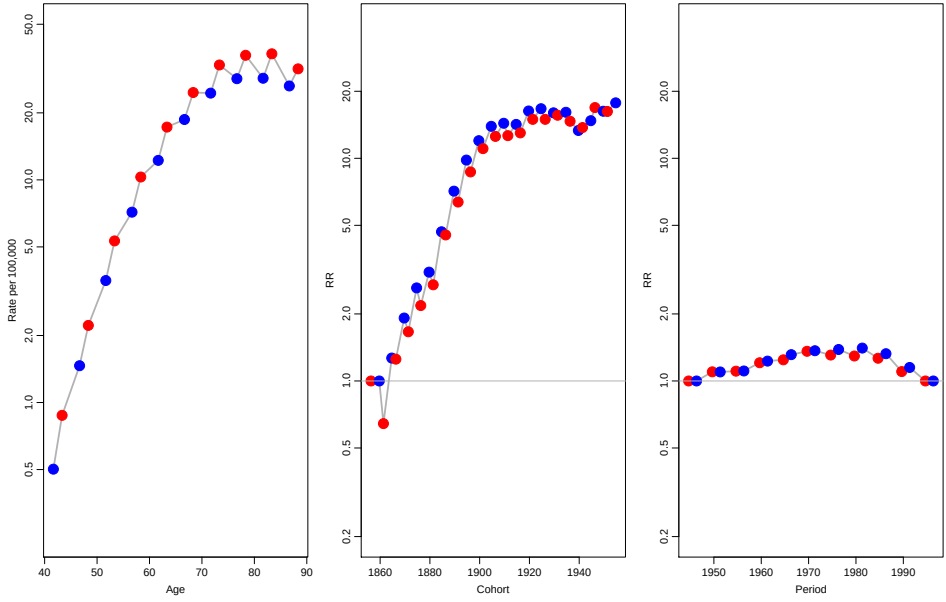
```
> library( Epi )
> data( lungDK )
> lungDK[1:10,]
   A5   P5   C5 up       Ax       Px       Cx  D        Y
1  40 1943 1898  1 43.33333 1944.667 1901.333 52 336233.8
2  40 1943 1903  0 41.66667 1946.333 1904.667 28 357812.7
3  40 1948 1903  1 43.33333 1949.667 1906.333 51 363783.7
4  40 1948 1908  0 41.66667 1951.333 1909.667 30 390985.8
5  40 1953 1908  1 43.33333 1954.667 1911.333 50 391925.3
6  40 1953 1913  0 41.66667 1956.333 1914.667 23 377515.3
7  40 1958 1913  1 43.33333 1959.667 1916.333 56 365575.5
8  40 1958 1918  0 41.66667 1961.333 1919.667 43 383689.0
9  40 1963 1918  1 43.33333 1964.667 1921.333 44 385878.5
10 40 1963 1923  0 41.66667 1966.333 1924.667 38 371361.5
```


Rate per 100,000
200
100
50
20
10
5
2
1
40
50
60
70
80
90
Age
RR
20.0
10.0
5.0
2.0
1.0
0.5
0.2
0.1
1860
1880
1900
1920
1940
Cohort
RR
20.0
10.0
5.0
2.0
1.0
0.5
0.2
0.1
1950
1960
1970
1980
1990
Period


Rate per 100,000
200
100
50
20
10
5
2
1
40
50
60
70
80
90
Age
RR
20.0
10.0
5.0
2.0
1.0
0.5
0.2
0.1
1860
1880
1900
1920
1940
Cohort
RR
20.0
10.0
5.0
2.0
1.0
0.5
0.2
0.1
1950
1960
1970
1980
1990
Period

Now, separately fit models for upper and lower triangles:

```
> mx.u <- glm( D ~ factor(Ax) - 1 +
+                 factor(Cx) +
+                 factor(Px) + offset( log( Y/10^5 ) ), family=poisson,
+                 data=lungDK[lungDK$up==1,] )
> mx.l <- glm( D ~ factor(Ax) - 1 +
+                 factor(Cx) +
+                 factor(Px) + offset( log( Y/10^5 ) ), family=poisson,
+                 data=lungDK[lungDK$up==0,] )
> mx$deviance
[1] 284.7269
> mx.l$deviance
[1] 134.4566
> mx.u$deviance
[1] 150.2703
> mx.l$deviance+mx.u$deviance
[1] 284.7269
```


Rate per 100,000
0.5
1.0
2.0
5.0
10.0
20.0
50.0
40
50
60
70
80
90
Age
RR
0.2
0.5
1.0
2.0
5.0
10.0
20.0
1860
1880
1900
1920
1940
Cohort
RR
0.2
0.5
1.0
2.0
5.0
10.0
20.0
1950
1960
1970
1980
1990
Period

What's the problem?

- One parameter is assigned to each distinct value of the timescales:

What's the problem?

- One parameter is assigned to each distinct value of the timescales:
- ordering and position of the observations on the scales is not used in the modelling

What's the problem?

- One parameter is assigned to each distinct value of the timescales:
- ordering and position of the observations on the scales is not used in the modelling
- — the exchangeability assumption again

What's the problem?

- One parameter is assigned to each distinct value of the timescales:
- ordering and position of the observations on the scales is not used in the modelling
- — the exchangeability assumption again
- Solution: model effects with smooth functions of the **mean** age, period and cohort with three functions:

$$\lambda_{ap} = f(a) + g(p) + h(c)$$

What's the problem?

- One parameter is assigned to each distinct value of the timescales:
- ordering and position of the observations on the scales is not used in the modelling
- — the exchangeability assumption again
- Solution: model effects with smooth functions of the **mean** age, period and cohort with three functions:

$$\lambda_{ap} = f(a) + g(p) + h(c)$$

- Fixes the problem with non-equidistant age, period and cohort classes

What's the problem?

- One parameter is assigned to each distinct value of the timescales:
- ordering and position of the observations on the scales is not used in the modelling
- — the exchangeability assumption again
- Solution: model effects with smooth functions of the **mean** age, period and cohort with three functions:

$$\lambda_{ap} = f(a) + g(p) + h(c)$$

- Fixes the problem with non-equidistant age, period and cohort classes
- The practical problem is how to choose a reasonable parametrization of these functions, and how to get estimates,

The identifiability problem still exists:

$$c = p - a \quad \Leftrightarrow \quad p - a - c = 0$$

The identifiability problem still exists:

$$c = p - a \quad \Leftrightarrow \quad p - a - c = 0$$

$$\begin{aligned}
\lambda_{ap} &= f(a) + g(p) + h(c) \\
&= f(a) + g(p) + h(c) + \gamma(p - a - c) \\
&= f(a) - \mu_a - \gamma a + \\
&\quad g(p) + \mu_a + \mu_c + \gamma p + \\
&\quad h(c) - \mu_c - \gamma c
\end{aligned}$$

The identifiability problem still exists:

$$c = p - a \quad \Leftrightarrow \quad p - a - c = 0$$

$$\begin{aligned}
\lambda_{ap} &= f(a) + g(p) + h(c) \\
&= f(a) + g(p) + h(c) + \gamma(p - a - c) \\
&= f(a) - \mu_a - \gamma a + \\
&\quad g(p) + \mu_a + \mu_c + \gamma p + \\
&\quad h(c) - \mu_c - \gamma c
\end{aligned}$$

A decision on parametrization is needed.
It must be **external to the model**.

The identifiability problem still exists:

$$c = p - a \quad \Leftrightarrow \quad p - a - c = 0$$

$$\begin{aligned}
\lambda_{ap} &= f(a) + g(p) + h(c) \\
&= f(a) + g(p) + h(c) + \gamma(p - a - c) \\
&= f(a) - \mu_a - \gamma a + \\
&\quad g(p) + \mu_a + \mu_c + \gamma p + \\
&\quad h(c) - \mu_c - \gamma c
\end{aligned}$$

A decision on parametrization is needed.
It must be **external to the model**.
. . . and is alien to the chosen parametrization of the APC-effects

Parametrization principle

1. The age-function should be interpretable as log age-specific rates in cohort c_0 after adjustment for the period effect.

Parametrization principle

1. The age-function should be interpretable as log age-specific rates in cohort c_0 after adjustment for the period effect.
2. The cohort function is 0 at a reference cohort c_0, interpretable as log-RR relative to cohort c_0.

Parametrization principle

1. The age-function should be interpretable as log age-specific rates in cohort c_0 after adjustment for the period effect.
2. The cohort function is 0 at a reference cohort c_0, interpretable as log-RR relative to cohort c_0.
3. The period function is 0 on average with 0 slope, interpretable as log-RR relative to the age-cohort prediction. (residual log-RR).

Implementation:

1. Obtain any set of parameters $\hat{f}(a)$, $\hat{g}(p)$, $\hat{h}(c)$.

Implementation:

1. Obtain any set of parameters $\hat{f}(a)$, $\hat{g}(p)$, $\hat{h}(c)$.
2. Extract the trend from the period effect:

$$\tilde{g}(p) = \hat{g}(p) - (\mu + \beta p)$$

Implementation:

1. Obtain any set of parameters $\hat{f}(a)$, $\hat{g}(p)$, $\hat{h}(c)$.
2. Extract the trend from the period effect:

$$\tilde{g}(p) = \hat{g}(p) - (\mu + \beta p)$$

3. Then use the functions:

$$\begin{aligned}
\tilde{f}(a) &= \hat{f}(a) + \mu + \beta a + \hat{h}(c_0) + \beta c_0 \\
\tilde{g}(p) &= \hat{g}(p) - \mu - \beta p \\
\tilde{h}(c) &= \hat{h}(c) \qquad + \beta c - \hat{h}(c_0) - \beta c_0
\end{aligned}$$

Implementation:

1. Obtain any set of parameters $\hat{f}(a)$, $\hat{g}(p)$, $\hat{h}(c)$.
2. Extract the trend from the period effect:

$$\tilde{g}(p) = \hat{g}(p) - (\mu + \beta p)$$

3. Then use the functions:

$$\begin{aligned}
\tilde{f}(a) &= \hat{f}(a) + \mu + \beta a + \hat{h}(c_0) + \beta c_0 \\
\tilde{g}(p) &= \hat{g}(p) - \mu - \beta p \\
\tilde{h}(c) &= \hat{h}(c) \qquad + \beta c - \hat{h}(c_0) - \beta c_0
\end{aligned}$$

4. Extracting trend requires an **inner product** to project colums of $g(p)$ on the orthogonal of $(1 \vdots p)$

Implementation:

1. Obtain any set of parameters $\hat{f}(a)$, $\hat{g}(p)$, $\hat{h}(c)$.
2. Extract the trend from the period effect:

$$\tilde{g}(p) = \hat{g}(p) - (\mu + \beta p)$$

3. Then use the functions:

$$\begin{aligned}
\tilde{f}(a) &= \hat{f}(a) + \mu + \beta a + \hat{h}(c_0) + \beta c_0 \\
\tilde{g}(p) &= \hat{g}(p) - \mu - \beta p \\
\tilde{h}(c) &= \hat{h}(c) \qquad + \beta c - \hat{h}(c_0) - \beta c_0
\end{aligned}$$

4. Extracting trend requires an **inner product** to project colums of $g(p)$ on the orthogonal of $(1 \vdots p)$

Implementation:

1. Obtain any set of parameters $\hat{f}(a)$, $\hat{g}(p)$, $\hat{h}(c)$.
2. Extract the trend from the period effect:

$$\tilde{g}(p) = \hat{g}(p) - (\mu + \beta p)$$

3. Then use the functions:

$$\begin{aligned}
\tilde{f}(a) &= \hat{f}(a) + \mu + \beta a + \hat{h}(c_0) + \beta c_0 \\
\tilde{g}(p) &= \hat{g}(p) - \mu - \beta p \\
\tilde{h}(c) &= \hat{h}(c) \qquad + \beta c - \hat{h}(c_0) - \beta c_0
\end{aligned}$$

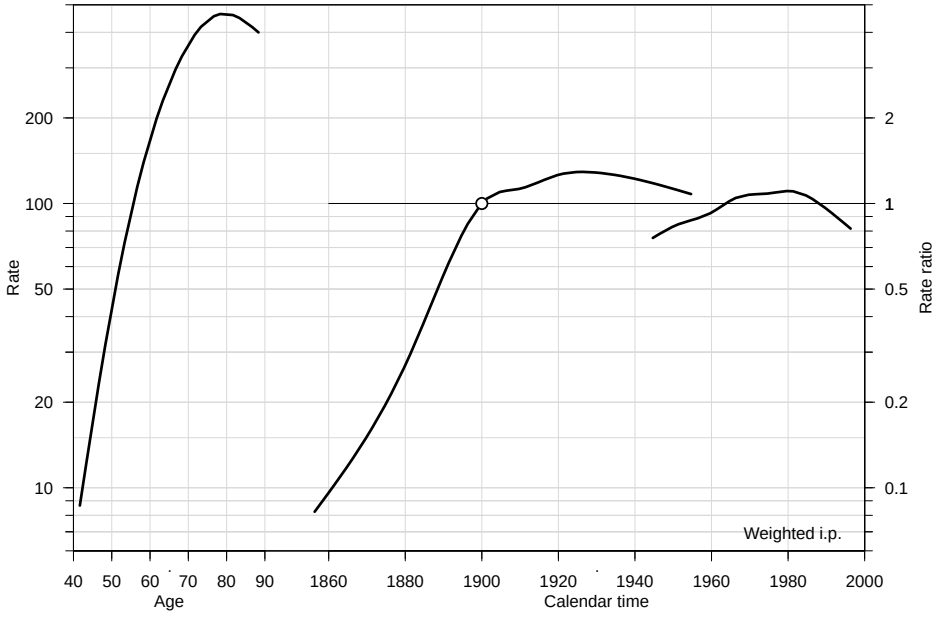
4. Extracting trend requires an **inner product** to project colums of $g(p)$ on the orthogonal of $(1 \vdots p)$, in the literature implicitly assumed to be induced by the identity, — a bold assumption

How to?

Implemented in `apc.fit` in the `Epi` package

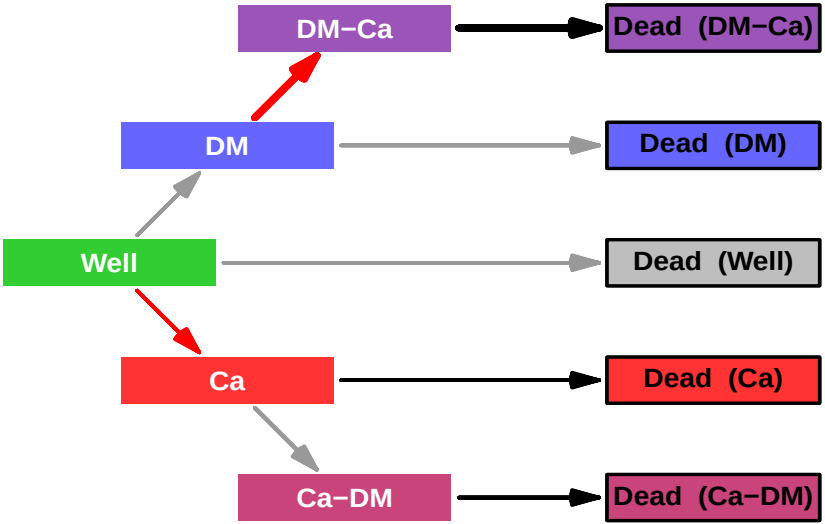
```
m1 <- apc.fit( A=lungDK$Ax,
               P=lungDK$Px,
               D=lungDK$D,
               Y=lungDK$Y/10^5,
           ref.c=1900 )
apc.plot( m1 )
```

Consult the help page for details.

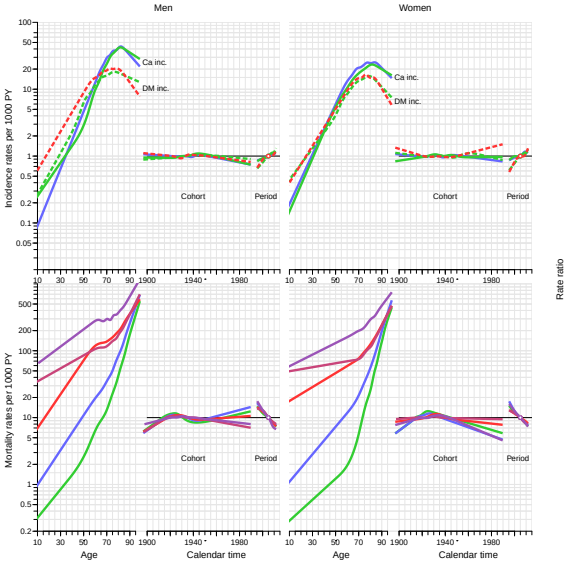

Rate
200
100
50
20
10
Rate ratio
2
1
0.5
0.2
0.1
Weighted i.p.
40
50
60
70
80
90
Age
1860
1880
1900
1920
1940
1960
1980
2000
Calendar time

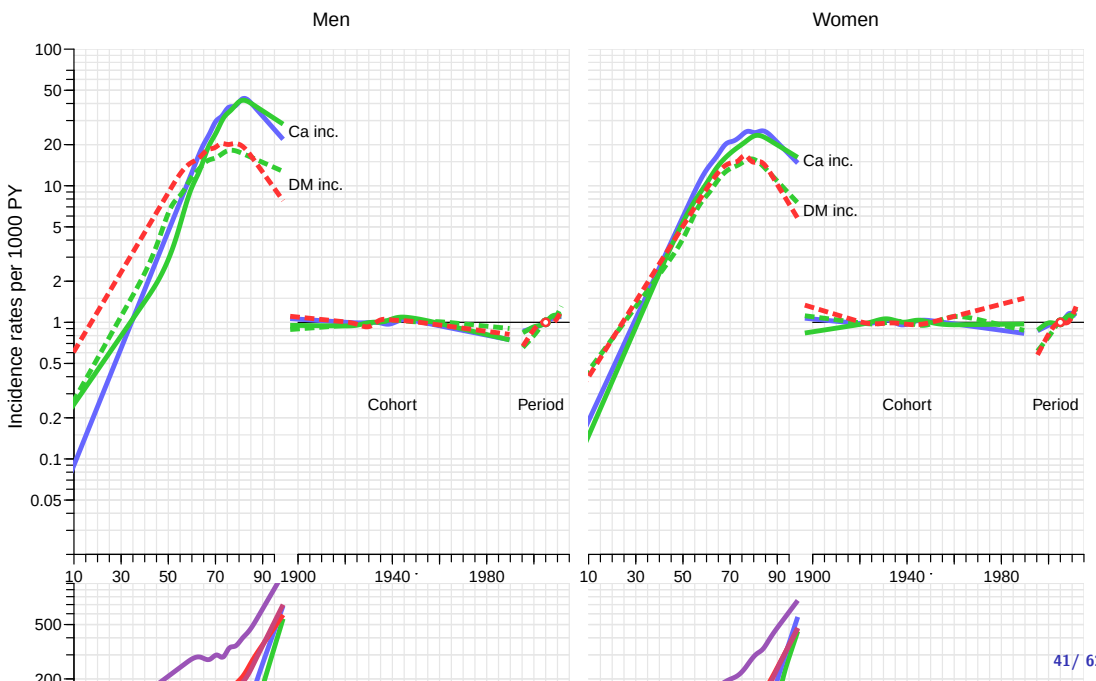

Rate
200
100
50
20
10
Rate ratio
2
1
0.5
0.2
0.1
Naïve i.p.
Weighted i.p.
40
50
60
70
80
90
Age
1860
1880
1900
1920
1940
1960
1980
2000
Calendar time

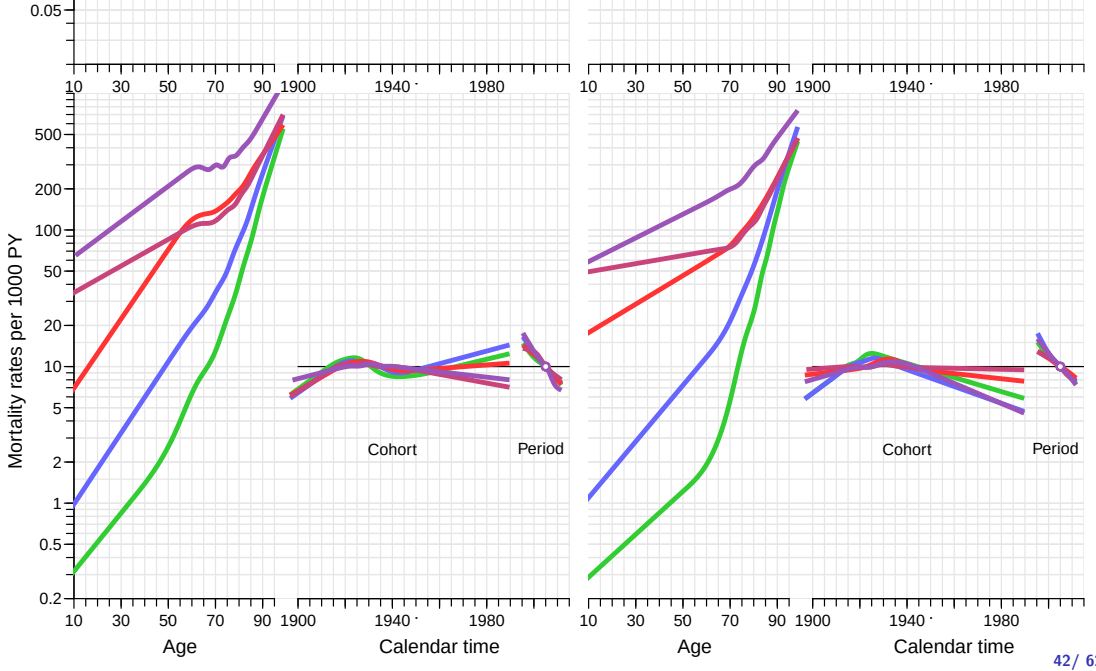
Joint occurrence of Diabetes and Cancer



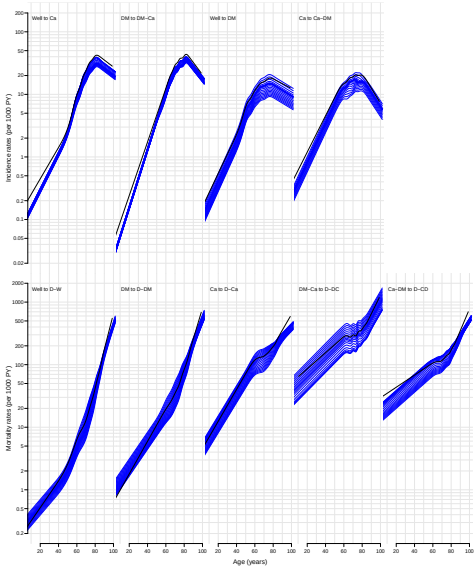
Joint occurrence of Diabetes and Cancer

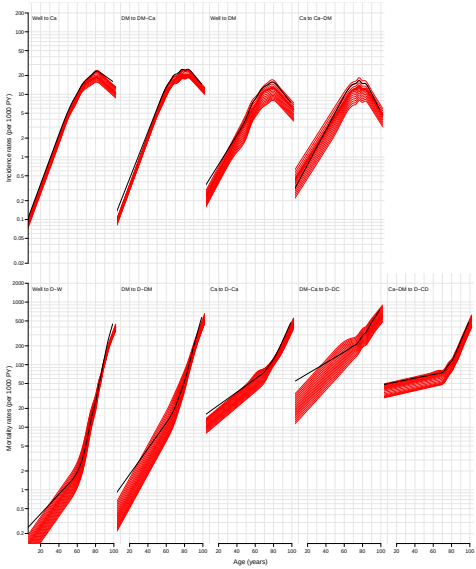



Men
Women
Incidence rates per 1000 PY
100
50
20
10
5
2
1
0.5
0.2
0.1
0.05
Ca inc.
DM inc.
Cohort
Period
10
30
50
70
90
1900
1940
1980
500
200


Mortality rates per 1000 PY
0.05
500
200
100
50
20
10
5
2
1
0.5
0.2
10
30
50
70
90
1900
1940
1980
Age
Calendar time
Cohort
Period

Predicted rates — cross-sectional 1995–2010





Continuous rates

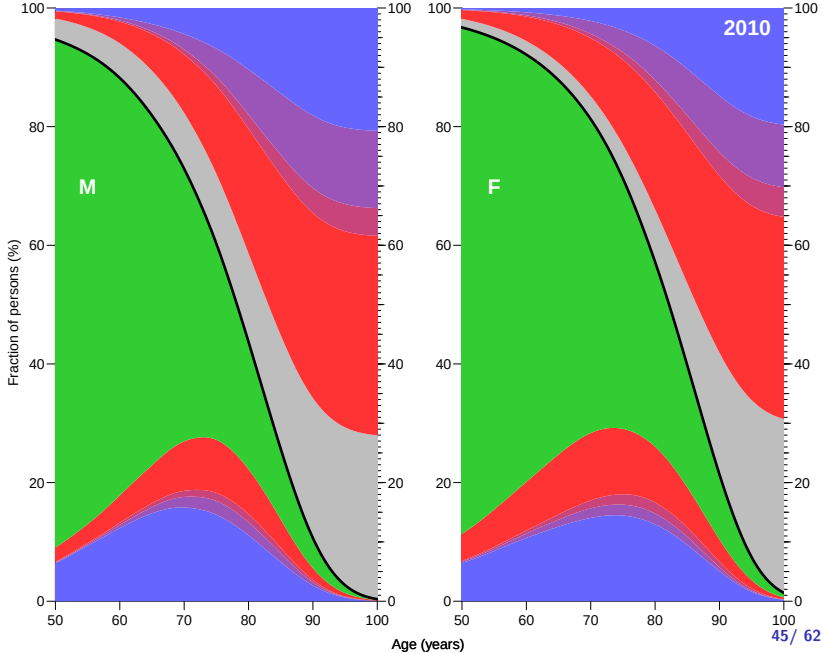
1-month cumulative rates $\rightarrow$ transition probabilities

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

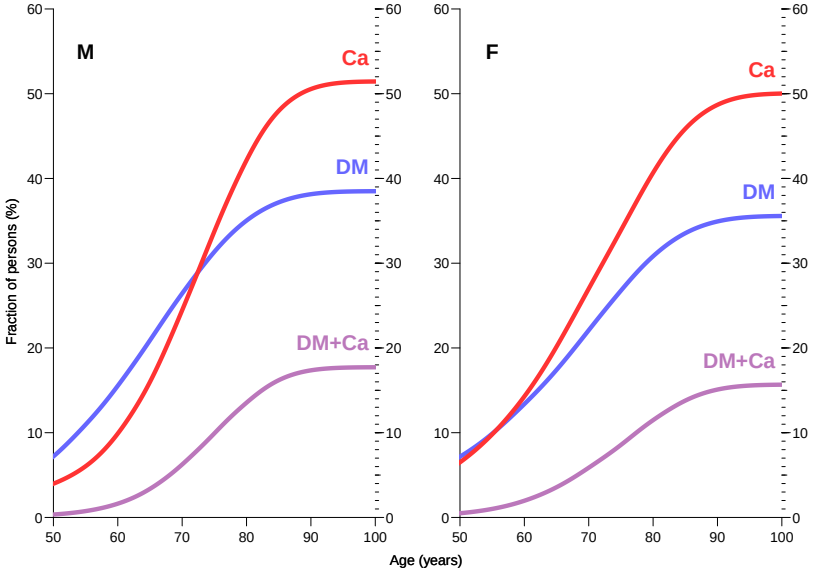
1-month transition probabilities ($\times 10^4$) at age 66 years:

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

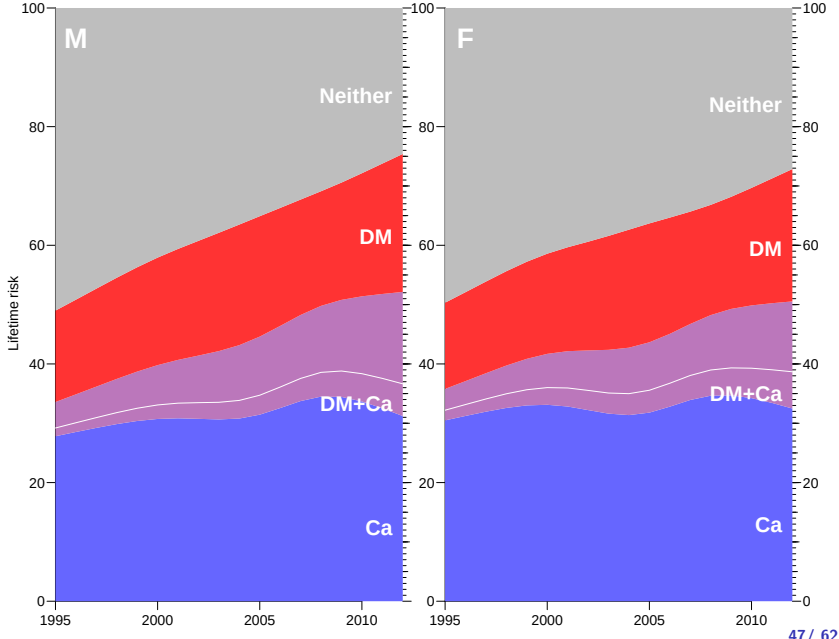
State occupancy probabilities



Lifetime risk



Trend in lifetime risk



Continuous time rates

- Transition rates between states:

Continuous time rates

- Transition rates between states:
  - based on 1-year tabulation of data

Continuous time rates

- Transition rates between states:
  - based on 1-year tabulation of data
  - age-period-cohort models

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

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
- State probabilities simple closed-form function of rates

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
- State probabilities simple closed-form function of rates
- Numerical integration of known functions trivial

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
- State probabilities simple closed-form function of rates
- Numerical integration of known functions trivial
- Matrix multiplication trivial

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
- State probabilities simple closed-form function of rates
- Numerical integration of known functions trivial
- Matrix multiplication trivial

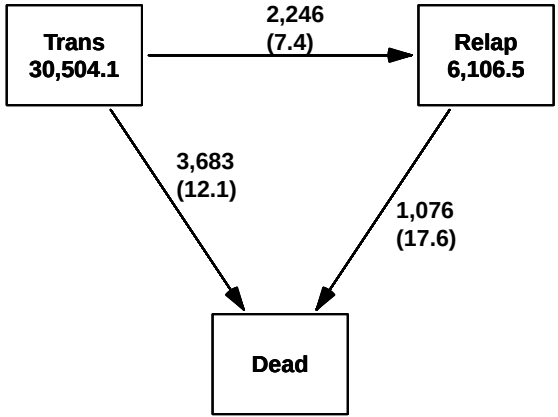
Continuous time rates

- Transition rates between states:
  - based on 1-year tabulation of data
  - age-period-cohort models
  - using smooth effects of age, period and cohort
- State probabilities simple closed-form function of rates
- Numerical integration of known functions trivial
- Matrix multiplication trivial

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

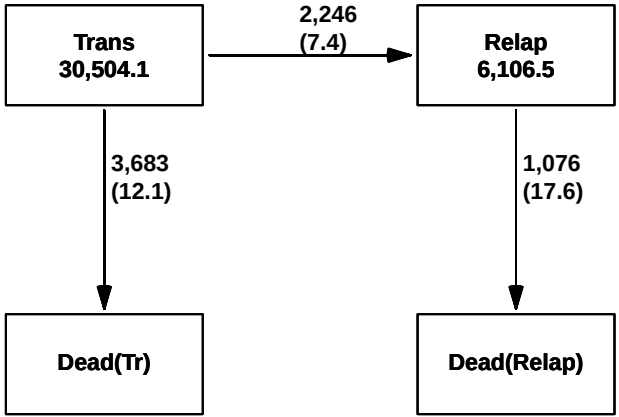
EBMT transplant data

Iacobelli & Carstensen: Multistate Models with Multiple Timescales, Stat Med 2013, [5]



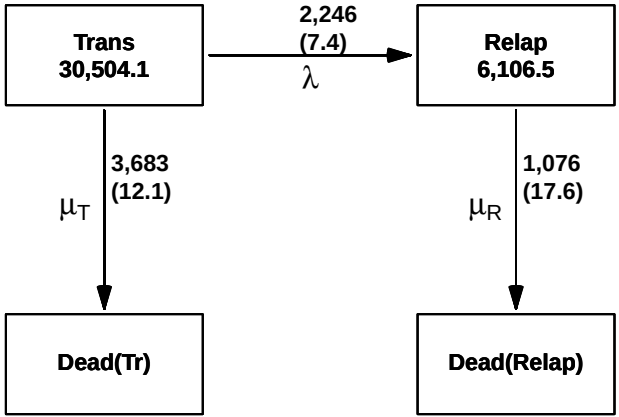
EBMT transplant data

Iacobelli & Carstensen: Multistate Models with Multiple Timescales, Stat Med 2013, [5]



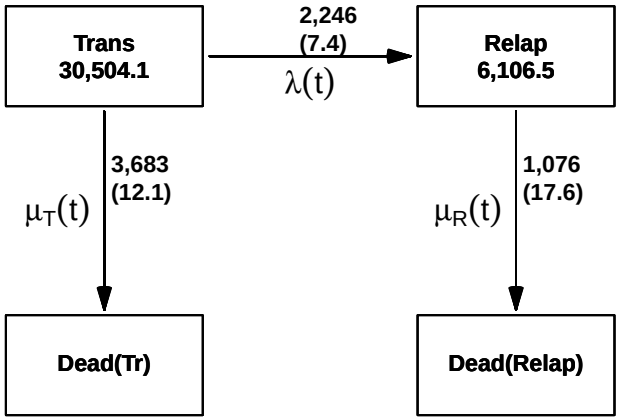
EBMT transplant data

Iacobelli & Carstensen: Multistate Models with Multiple Timescales, Stat Med 2013, [5]



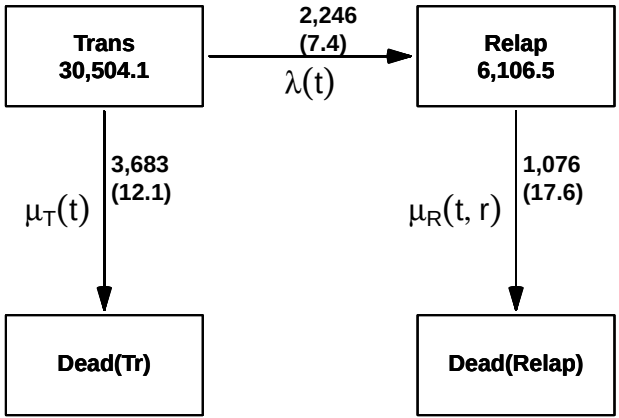
EBMT transplant data

Iacobelli & Carstensen: Multistate Models with Multiple Timescales, Stat Med 2013, [5]



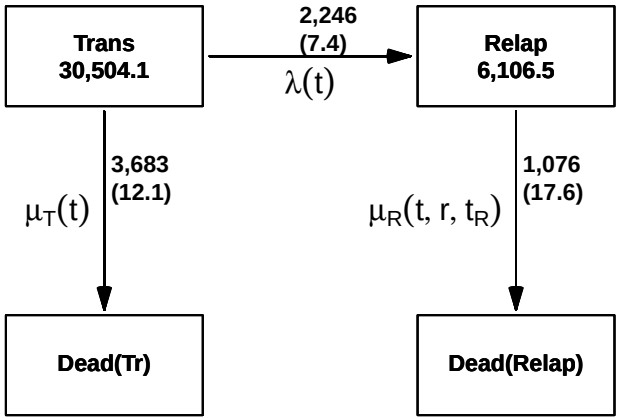
EBMT transplant data

Iacobelli & Carstensen: Multistate Models with Multiple Timescales, Stat Med 2013, [5]



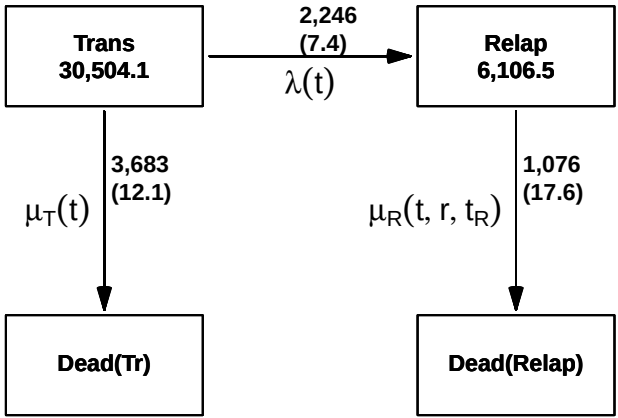
EBMT transplant data

Iacobelli & Carstensen: Multistate Models with Multiple Timescales, Stat Med 2013, [5]

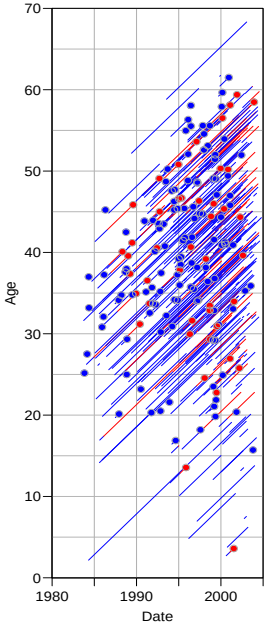


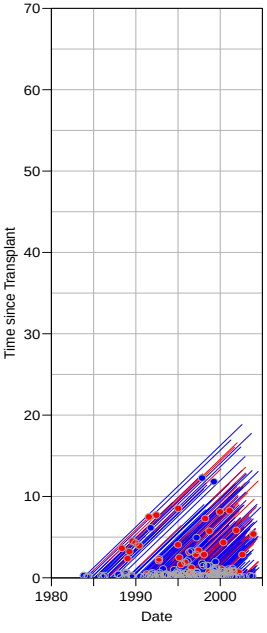
EBMT transplant data

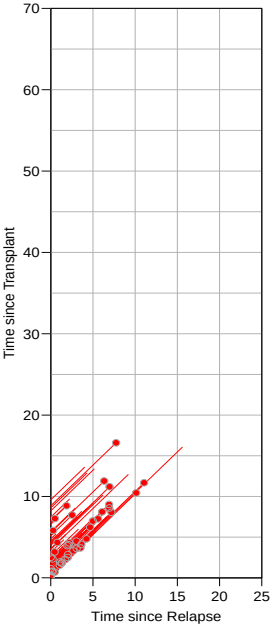
Iacobelli & Carstensen: Multistate Models with Multiple Timescales, Stat Med 2013, [5]



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


Age
Date
0
10
20
30
40
50
60
70
1980
1990
2000


Time since Transplant
Date
0
10
20
30
40
50
60
70
1980
1990
2000


Time since Transplant
Time since Relapse
0
10
20
30
40
50
60
70
0
5
10
15
20
25

Markov property: Empirical question

Model for mortality rates:

- t time since transplant

Markov property: Empirical question

Model for mortality rates:

- t time since transplant
- r time since relapse (if relapsed)

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

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:

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

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

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.

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**

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 [6, 7]

```
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  ...
```

Using the Lexis machinery [6, 7]

```
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  ...
```

Using the Lexis machinery [6, 7]

```
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  ...
```

Using the Lexis machinery [6, 7]

```
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   ...
```

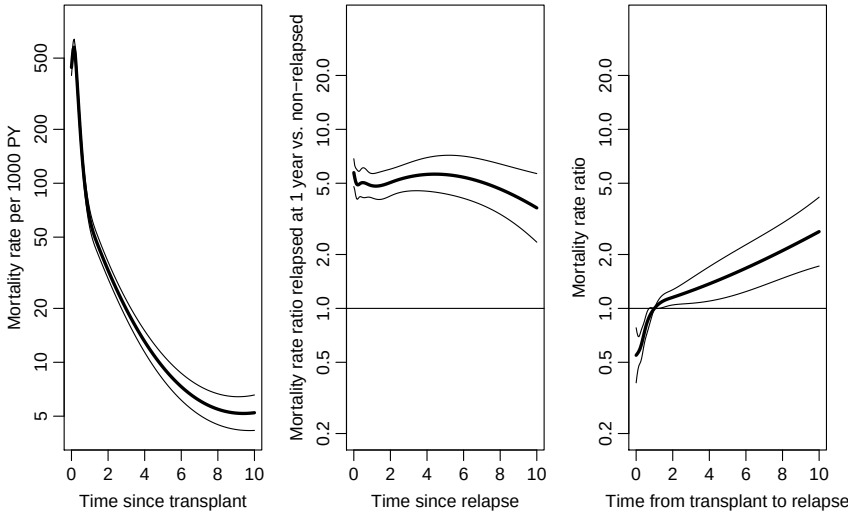
Using the Lexis machinery [6, 7]

```
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  ...
```

Using the Lexis machinery [6, 7]

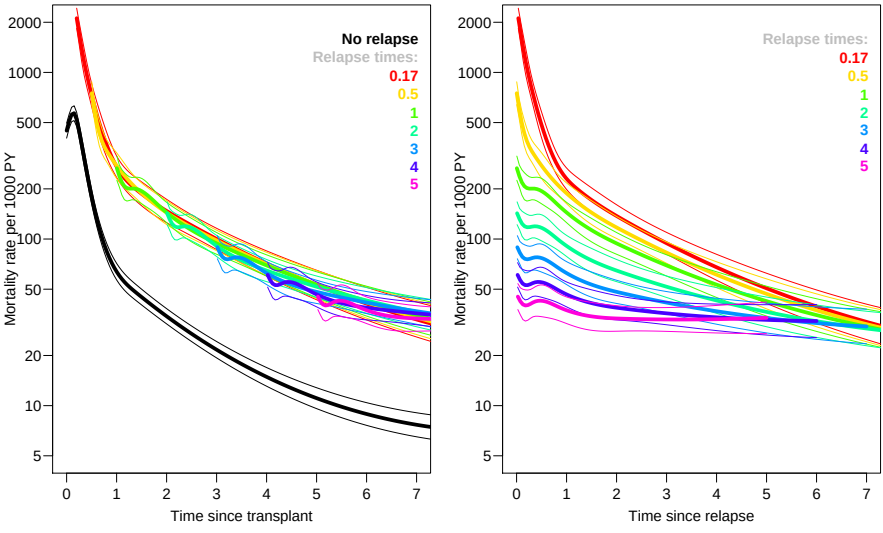
```
cmlT <- Lexis(entry = list(cal = cal.yr(dot),
                           age = cal.yr(dot)-cal.yr(dob),
                           tst = 0),
              exit = list(cal = cal.yr(dof)),
       exit.status = dead,
            states = c("Transplant","Dead"),
              data = cml )

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   ...
```

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



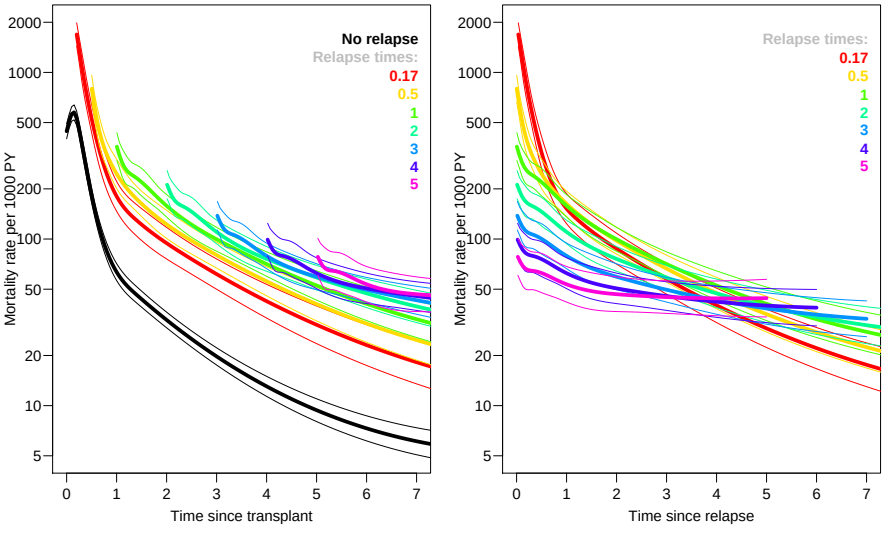
t: time since transplant r: time since relapse

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



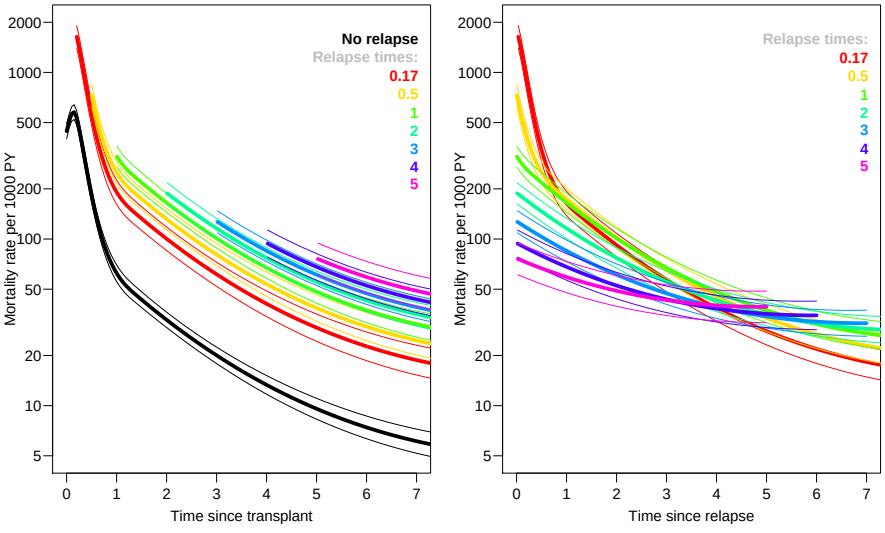
t: time since transplant r: time since relapse

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



t: time since transplant r: time since relapse

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



t: time since transplant $\quad$ r: time since relapse

Model summary

- Mortality of relapsed patients depends on **when** they relapsed.

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.

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.

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)

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

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

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

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

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

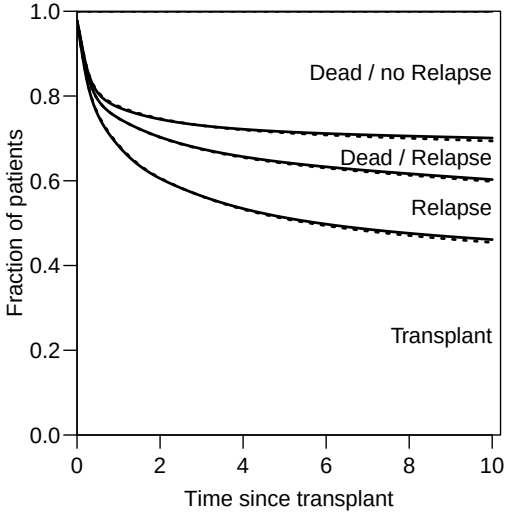
Not Markov: the hard way

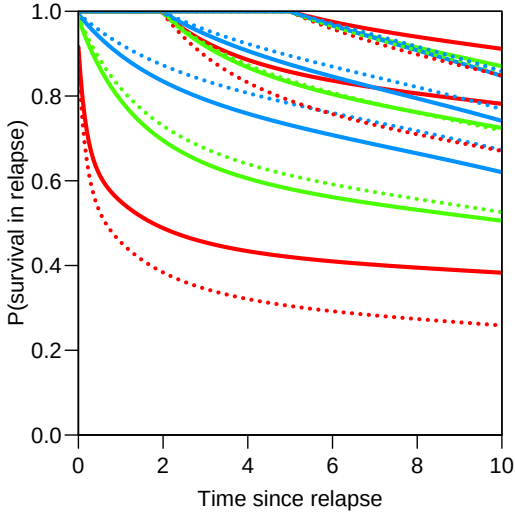
$$\mathrm{P}\{\mathsf{T\ at}\ t\} = \exp\Big(-\int_0^t \lambda(s) + \mu_T(s)\,\mathrm{d}s\Big)$$

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

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

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





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

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying
- Continuous time formulae easiest to handle

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying
- Continuous time formulae easiest to handle
- Statistical models should reflect this:

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying
- Continuous time formulae easiest to handle
- Statistical models should reflect this:
  - Parametric form of time-effects allow direct implementation of probability theory

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying
- Continuous time formulae easiest to handle
- Statistical models should reflect this:
  - Parametric form of time-effects allow direct implementation of probability theory
  - Corrolary: Choice of time scales is an **empirical** problem

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying
- Continuous time formulae easiest to handle
- Statistical models should reflect this:
  - Parametric form of time-effects allow direct implementation of probability theory
  - Corrolary: Choice of time scales is an **empirical** problem
- Reporting of models should reflect this

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying
- Continuous time formulae easiest to handle
- Statistical models should reflect this:
  - Parametric form of time-effects allow direct implementation of probability theory
  - Corrolary: Choice of time scales is an **empirical** problem
- Reporting of models should reflect this
- Stick to this world:

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying
- Continuous time formulae easiest to handle
- Statistical models should reflect this:
  - Parametric form of time-effects allow direct implementation of probability theory
  - Corrolary: Choice of time scales is an **empirical** problem
- Reporting of models should reflect this
- Stick to this world:

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying
- Continuous time formulae easiest to handle
- Statistical models should reflect this:
  - Parametric form of time-effects allow direct implementation of probability theory
  - Corrolary: Choice of time scales is an **empirical** problem
- Reporting of models should reflect this
- Stick to this world: Fewer tables — more graphs!

Summary & Conclusions

- The world is continuous
- Effects of time likely to be continuously, smoothly varying
- Continuous time formulae easiest to handle
- Statistical models should reflect this:
  - Parametric form of time-effects allow direct implementation of probability theory
  - Corrolary: Choice of time scales is an **empirical** problem
- Reporting of models should reflect 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, Nov 2011.

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

X. Yang, G. T. Ko, W. Y. So, R. C. Ma, L. W. Yu, A. P. Kong, H. Zhao, C. C. Chow, P. C. Tong, and J. C. Chan.
Associations of hyperglycemia and insulin usage with the risk of cancer in type 2 diabetes: the Hong Kong diabetes registry.
Diabetes, 59:1254–1260, May 2010.

B. Carstensen.
Comment on: Yang et al. (2010) Associations of hyperglycemia and insulin usage with the risk of cancer in type 2 diabetes: the Hong Kong Diabetes Registry. Diabetes;59:1254-1260.
Diabetes, 59:e17–18, Sep 2010.

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.