# Representation and prediction in multistate models

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

MPIDR, Berlin, Germany, 24–25 November 2016



## Representation of FU

- Many different ways of representing follow-up:
  - `msm`
  - `mstate`
  - `etm`
- some dedicated to a special model
- mostly on one timescale (the `st` machinery in Stata)
- generally derived from follow-up in a clinical trial with
  - entry at $0$
  - exit at $t$
  - risk time $t$



## MSM representation

Epochs (time-intervals)

- **Time** at start: The value of the timescale**s** of interest.
- **Length** of the epoch — risk time ("exposure")
- **State** in which the the time is spent
- **next** state: the state the person moves on to

This is a representation of the **risk time**

## Representation of likelihood

- The likelihood terms in a multistate model comes naturally as one per **transition**:
- $$\exp\left(\int_{\text{in}}^{\text{out}} \lambda_j(u)\,\mathrm{d}u\right)$$
- . . . and for the events, additionally $\lambda_j(\text{out})$
- This is the **stacked** representation, allowing joint modeling of all transition rates in one go:
- Separate or joint parameters for each transition.



## `Lexis` objects

`data.frame` with a few extra features:

timescales — chose any number, name as you please

`lex.id` — person identification

`lex.dur` — length of the epoch represented

`lex.Cst` — state of risk time (Current `st`ate)

`lex.Xst` — ne`X`t `st`ate



## `Lexis` objects - split objects

Splitting time with `splitLexis`:

- risk time in many small intervals
- assuming constant rates in each
- allows simple parametric modeling with `glm`
- using any number of time scales
- prediction of rates
- calculation of any derived measure

## glm modeling form object Lx

```
> w2i <- glm( lex.Xst=="Ill" ~ v1 + v2 + v3,
+             family=poisson, offset=log(lex.dur),
+             data = subset( Lx, lex.Cst=="Well" ) )
> #
> w2d <- glm( lex.Xst=="Dead" ~ v1 + v2 + v3,
+             family=poisson, offset=log(lex.dur),
+             data = subset( Lx, lex.Cst=="Well" ) )
> #
> i2d <- glm( lex.Xst=="Dead" ~ v1 + v2 + v3,
+             family=poisson, offset=log(lex.dur),
+             data = subset( Lx, lex.Cst=="Ill" ) )
```



## coxph modeling form object Lx

```
> w2i <- coxph( Surv( time, time+lex.dur, lex.Xst=="Ill" ) ~ v1 + v2 + v3,
+               data = subset( Lx, lex.Cst=="Well" ) )
> #
> w2d <- coxph( Surv( time, time+lex.dur, lex.Xst=="Dead" ) ~ v1 + v2 + v3,
+               data = subset( Lx, lex.Cst=="Well" ) )
> #
> i2d <- coxph( Surv( time, time+lex.dur, lex.Xst=="Dead" ) ~ v1 + v2 + v3,
+               data = subset( Lx, lex.Cst=="Ill" ) )
```



## glm modeling form object Lx with update

```
> w2i <- glm( lex.Xst=="Ill" ~ v1 + v2 + v3,
+             family=poisson, offset=log(lex.dur),
+             data = subset( Lx, lex.Cst=="Well" ) )
> #
> w2d <- update( w2i, lex.Xst=="Dead" ~ . )
> #
> i2d <- update( w2d, data = subset( Lx, lex.Cst=="Ill" ) )
```

— but if you really **are** using the same covariates:

```
> mA <- glm( lex.Fail ~ lex.Tr + (v1 + v2 + v3):lex.Tr,
+            family=poisson, offset=log(lex.dur),
+            data = stack.Lexis( Lx ) )
```

`lex.Tr` has levels "`Well->Ill`", "`Well->Dead`" and "`Ill->Dead`"

## Stacking – interface to `mstate`

- Stacking of the follow-up gives the terms that make up the likelihood
- `lex.Tr` has levels "`Well->Ill`", "`Well->Dead`" and "`Ill->Dead`"
- `lex.Fail` is logical.
- Used for the interface function `msdata.Lexis` — stacks the data and shaves away features.



## Utilities — summaries

- `summary` — gives table of transitions and risk time
- `boxes` — shows it in a nice transition diagram (boxes)
- `plot` — plots follow-up in a Lexis diagram



## Utilities — manipulation & restructuring

- `cutLexis` — cuts follow-up at a time point; transition to a new state.
- `splitLexis` — splits time in small intervals along a time scale at designated breaks.
- Both features need `stsplit` in Stata (+plus a bit of fidgeting)

## Predicting using a single time scale

- With a parametric model for rates:
- closed form expressions for state probabilities etc.
- wrapping in a function with parameters as arguments
- parametric bootstrap does all for you
- `multistate` machinery from Michael C



## Prediction using realistic models

- Simulation of an entire dataset through MS model
  - allowing multiple time scale
  - . . . also "time since state S"
- Enumeration of persons in states gives
- — state occupancy probabilities
- — expected sojourn times
- — expected life time probabilities



Diabetologia
DOI 10.1007/s00125-016-4065-6

CrossMark

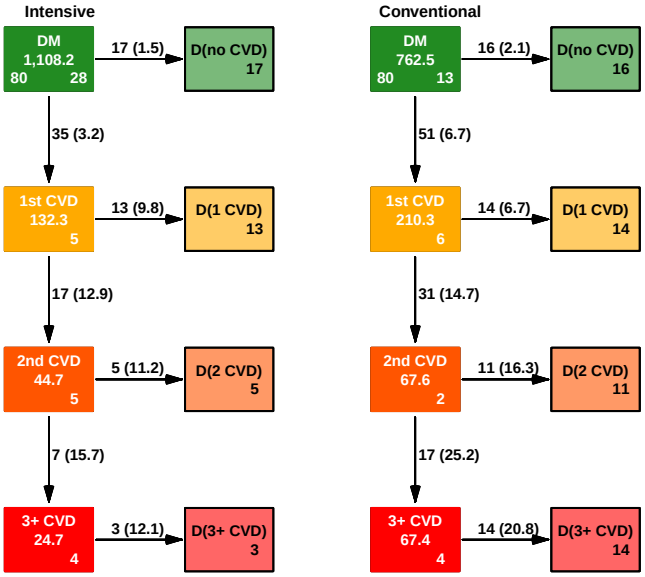
ARTICLE

# Years of life gained by multifactorial intervention in patients with type 2 diabetes mellitus and microalbuminuria: 21 years follow-up on the Steno-2 randomised trial

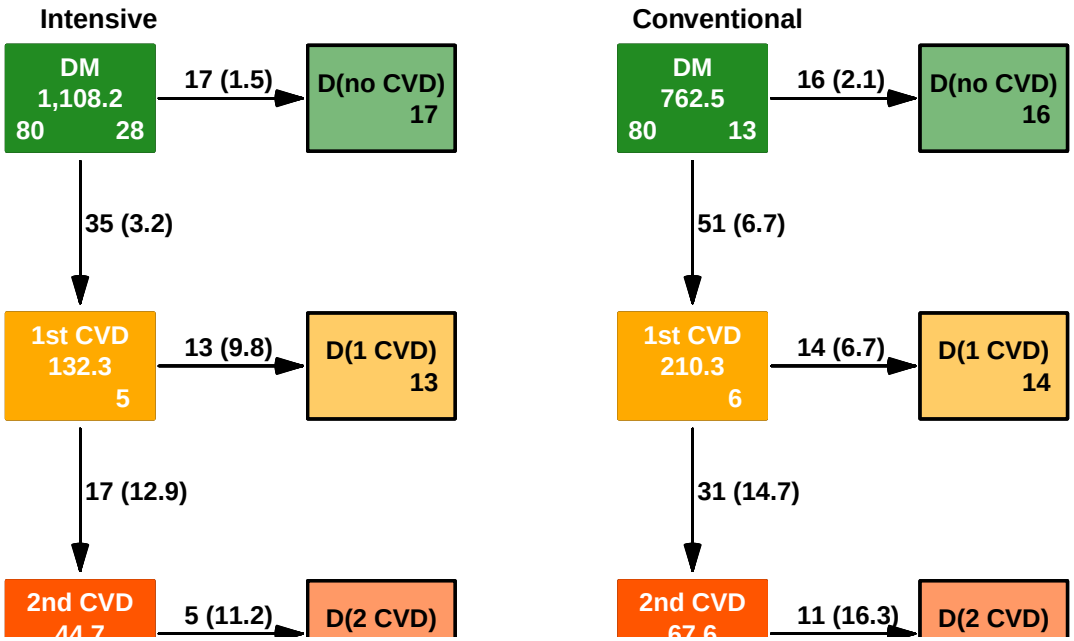
**Peter Gæde**[1,2] · **Jens Oellgaard**[1,2,3] · **Bendix Carstensen**[3] · **Peter Rossing**[3,4,5] · **Henrik Lund-Andersen**[3,5,6] · **Hans-Henrik Parving**[5,7] · **Oluf Pedersen**[8]

Received: 7 April 2016 / Accepted: 1 July 2016


**Abstract**
*Aims/hypothesis* The aim of this work was to study the potential long-term impact of a 7.8 years intensified, multifactorial pharmacological approaches. After 7.8 years the study continued as an observational follow-up with all patients receiving treatment as for the original intensive-therapy group. The pri-



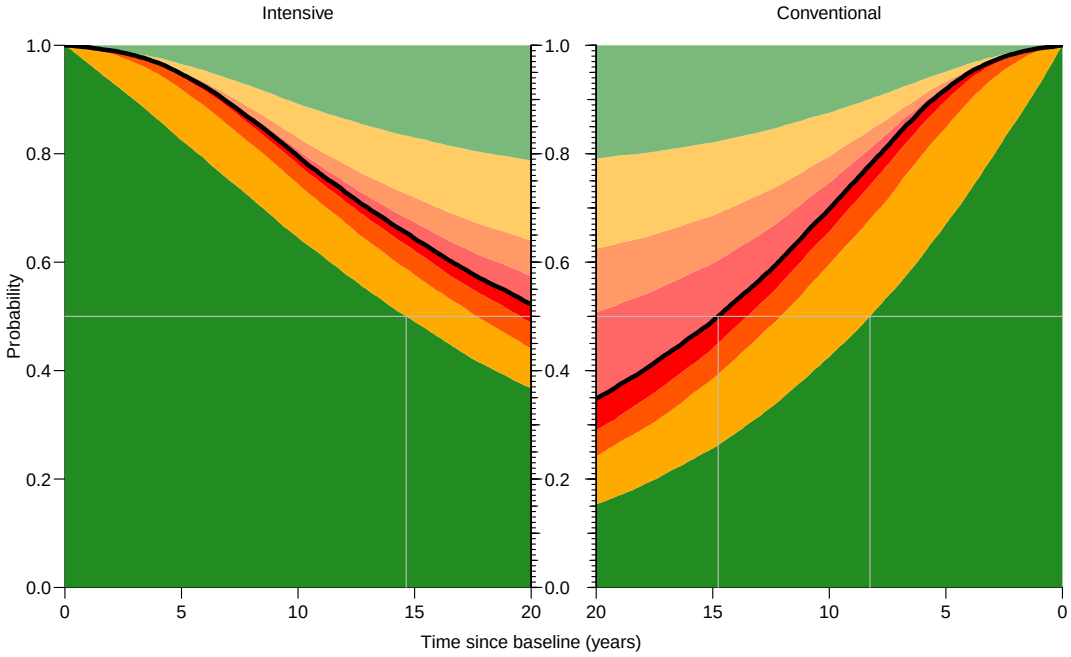


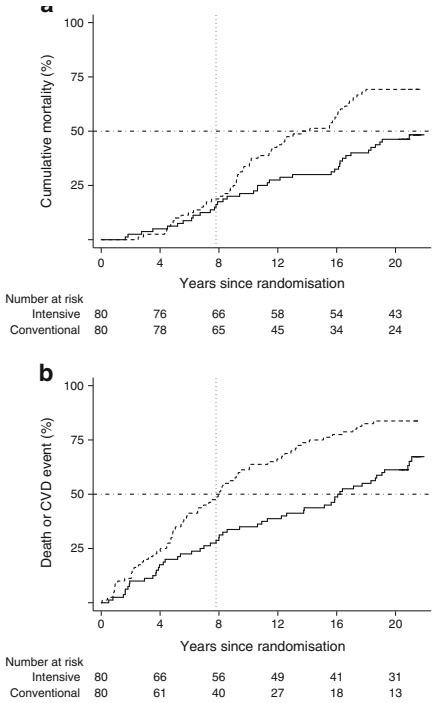



# Hazard ratios

| | CVD event | Mortality |
|---|---|---|
| HR, Int. vs. Conv. | 0.55 (0.39;0.77) | 0.83 (0.54; 1.30) |
| $H_0$: PH btw. CVD groups | p=0.438 | p=0.261 |
| $H_0$: HR = 1 | p=0.425 | p=0.001 |
| HR vs. 0 CVD events: | | |
| 0 (ref.) | 1.00 | 1.00 |
| 1 | 2.43 (1.67;3.52) | 3.08 (1.82; 5.19) |
| 2 | 3.48 (2.15;5.64) | 4.42 (2.36; 8.29) |
| 3+ | | 7.76 (4.11;14.65) |






between groups (HR 0.83 [95% CI 0.54, 1.30], $p=0.43$). Thus, the reduced mortality was primarily due to reduced risk of CVD.

The patients in the intensive group experienced a total of 90 cardiovascular events vs 195 events in the conventional group. Nineteen intensive-group patients (24%) vs 34 conventional-group patients (43%) experienced more than one cardiovascular event. No significant between-group difference in the distribution of specific cardiovascular first-event types was observed (Table 2 and Fig. 4).

**Microvascular complications** Hazard rates of progression rates in microvascular complications compared with baseline status are shown Fig. 3. Sensitivity analyses showed a negligible effect of the random dates imputation.

Progression of retinopathy was decreased by 33% in the intensive-therapy group (Fig. 5). Blindness in at least one eye was reduced in the intensive-therapy group with an HR of 0.47 (95% CI 0.23, 0.98, $p=0.044$). Autonomic neuropathy was decreased by 41% in the intensive-therapy group (Fig. 5). We observed no difference between groups in the progression of peripheral neuropathy (Fig. 5). Progression to diabetic nephropathy (macroalbuminuria) was reduced by 48% in the intensive-therapy group (Fig. 5). Ten patients in the conventional-therapy groups vs five patients in the intensive-therapy group progressed to end-stage renal disease ($p=0.061$).

**Discussion**



# Expected lifetime and YLL (well, gained)

Expected lifetime (years) in the Steno 2 cohort during the first 20 years after baseline by treatment group and CVD status.

| State | Intensive | Conventional | Int.−Conv. |
|---|---|---|---|
| Alive | 15.6 | 14.1 | 1.5 |
| No CVD | 12.7 | 10.0 | 2.6 |
| Any CVD | 3.0 | 4.1 | −1.1 |

Expected lifetime (years) during the first 20 years after baseline by sex, age, treatment group and CVD status.

| sex | | Men | | | Women | | |
|---|---|---|---|---|---|---|---|
| state | age | Int. | Conv. | Int.−Conv. | Int. | Conv. | Int.−Conv. |
| Alive | 45 | 18.5 | 17.5 | 1.0 | 19.1 | 18.4 | 0.7 |
| | 50 | 17.2 | 16.1 | 1.1 | 18.0 | 17.2 | 0.8 |
| | 55 | 15.6 | 13.8 | 1.8 | 17.4 | 15.9 | 1.6 |
| | 60 | 13.9 | 11.6 | 2.2 | 15.5 | 13.7 | 1.8 |
| | 65 | 11.2 | 9.5 | 1.8 | 13.3 | 11.4 | 2.0 |
| No CVD | 45 | 14.9 | 12.5 | 2.4 | 15.8 | 14.3 | 1.5 |
| | 50 | 14.0 | 11.1 | 2.9 | 15.1 | 12.9 | 2.2 |
| | 55 | 12.2 | 9.7 | 2.5 | 14.3 | 11.6 | 2.7 |
| | 60 | 10.9 | 8.2 | 2.7 | 12.4 | 9.9 | 2.6 |
| | 65 | 9.0 | 6.7 | 2.2 | 10.7 | 8.3 | 2.4 |



## Multistate models in practice:

- Representation:
  - States
  - Transitions
  - Sojourn times
  - Rates
- Analysis of rates:
  - Cox-model
  - Poisson model
- Reporting
  - Rates
  - HRs
  - Probabilities
  - Expected lifetime



## Representation of multistate FU: Lexis

- Allowing multiple time scales
  - time-scale variables — the starting point on each time scale
  - sojourn time variable `lex.dur` — risktime, exposure
  - state variables:
- Allowing multiple states
  - `lex.Cst` — the state in which follow-up (`lex.dur`) occurs
  - `lex.Xst` — the state in which

```
> head( st2[,-c(1,4:5,12,13)] )

    allocation sex    doBth     doDM   doBase   doCVD1   doCVD2   doCVD3    doDth
1    Intensive   M 1932.300 1991.093 1993.357 2014.457       NA       NA       NA
2    Intensive   M 1946.813 1982.130 1993.357 2009.231 2009.570 2010.162       NA
3 Conventional   M 1943.377 1982.560 1993.362 2001.639       NA       NA 2001.639
4 Conventional   M 1944.771 1976.791 1993.362 1995.415 1997.461 2003.443 2003.443
5 Conventional   M 1935.993 1986.346 1993.162 1993.959 1994.657 1998.047 1998.047
6 Conventional   M 1946.942 1986.404 1993.201 1998.769       NA       NA       NA
```



```
> L2 <- Lexis( entry = list( per = doBase,
+                            age = doBase-doBth,
+                            dur = doBase-doDM,
+                            tsb = 0 ),
+              exit = list( per = doEnd ),
+              exit.status = factor( ( !is.na(doDth) ) + deathCVD,
+                                    labels=c("DM","D-oth","CV-D") ),
+              data = st2 )

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

> summary( L2 )

Transitions:
     To
From DM D-oth CV-D  Records:  Events: Risk time:  Persons:
  DM 67    55   38       160       93     2417.6       160
```



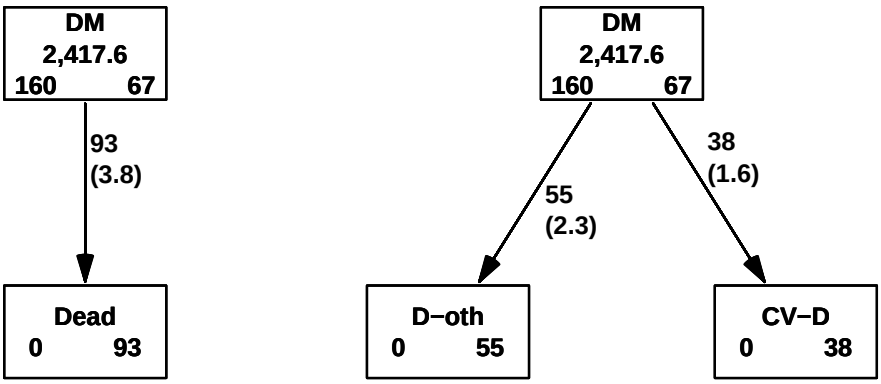
```
> L0 <- Lexis( entry = list( per = doBase,
+                            age = doBase-doBth,
+                            dur = doBase-doDM,
+                            tsb = 0 ),
+              exit = list( per = doEnd ),
+              exit.status = factor( ( !is.na(doDth) ),
+                                    labels=c("DM","Dead") ),
+              data = st2 )

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

> summary( L0 )

Transitions:
     To
From DM Dead  Records:  Events: Risk time:  Persons:
  DM 67   93       160       93     2417.6       160
```

```
> par( mfrow=c(1,2) )
> boxes( L0, boxpos=list(x=c(50,50),y=c(85,15)),show.BE=T,scale.R=100)
> boxes( L2, boxpos=list(x=c(50,15,85),y=c(85,15,15)),show.BE=T,scale.R=100)
```





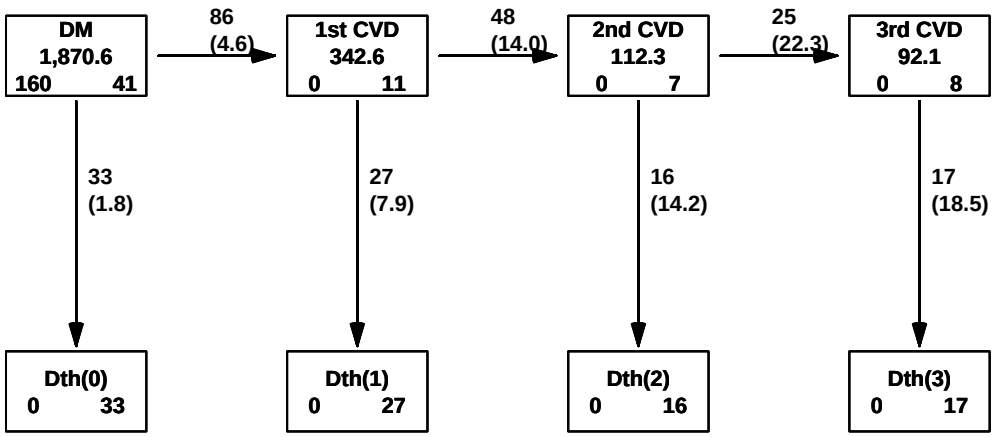

```
> L1 <- cutLexis( L0, cut = L0$doCVD1,
+                 new.state = "1st CVD",
+         precursor.states = c("DM"),
+             split.states = TRUE,
+                 new.scale = "tfCVD")
> #
> L1 <- cutLexis( L1, cut = L1$doCVD2,
+                 new.state = "2nd CVD",
+         precursor.states = c("DM","1st CVD"),
+             split.states = TRUE )
> #
> L1 <- cutLexis( L1, cut = L1$doCVD3,
+                 new.state = "3rd CVD",
+         precursor.states = c("DM","1st CVD","2nd CVD"),
+             split.states = TRUE )
> levels( L1 )
```

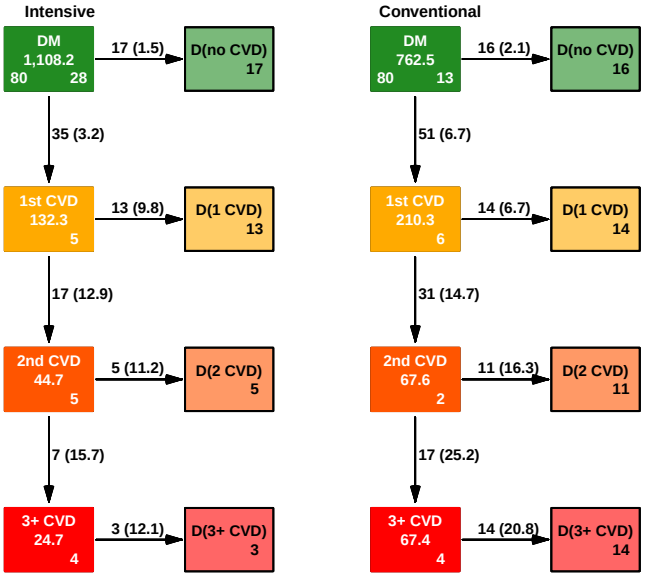
```
[1] "DM"                                "1st CVD"
[3] "2nd CVD"                           "3rd CVD"
[5] "Dead"                              "Dead(1st CVD)"
[7] "Dead(1st CVD)(2nd CVD)"            "Dead(1st CVD)(2nd CVD)(3rd CVD)"
```

```
> levels( L1$lex.Cst )[5:8] <-
+ levels( L1$lex.Xst )[5:8] <- paste("Dth(",0:3,")",sep="")
```



```
> boxes( L1, boxpos=list( x=c(rep(seq(15,85,,4),2) ),
+                         y=rep(c(85,20),each=4) ), show.BE=TRUE, scale.R=100 )
```








# An added time scale

We added the timescale `tfCVD` using `new.scale="tfCVD"`

```
> subset( L1, lex.id==5 )[,1:9]
```

| | per | age | dur | tsb | tfCVD | lex.dur | lex.Cst | lex.Xst | lex.id |
|---|---|---|---|---|---|---|---|---|---|
| 10 | 1993.162 | 57.16906 | 6.815916 | 0.0000000 | NA | 0.7967146 | DM | 1st CVD | 5 |
| 11 | 1993.959 | 57.96578 | 7.612631 | 0.7967146 | 0.000000 | 0.6981520 | 1st CVD | 2nd CVD | 5 |
| 12 | 1994.657 | 58.66393 | 8.310783 | 1.4948665 | 0.698152 | 3.3894593 | 2nd CVD | Dth(2) | 5 |



# An added time scale I

We added the timescale `tfCVD` using `new.scale="tfCVD"`

```
> S1 <- splitLexis( L1, breaks=seq(0,25,1/12), time.scale="tsb" )
> subset( S1, lex.id==5 )[,1:9]
```

| | lex.id | per | age | dur | tsb | tfCVD | lex.dur | lex.Cst | le |
|---|---|---|---|---|---|---|---|---|---|
| 741 | 5 | 1993.162 | 57.16906 | 6.815916 | 0.00000000 | NA | 0.08333333 | DM | |
| 742 | 5 | 1993.246 | 57.25240 | 6.899250 | 0.08333333 | NA | 0.08333333 | DM | |
| 743 | 5 | 1993.329 | 57.33573 | 6.982583 | 0.16666667 | NA | 0.08333333 | DM | |
| 744 | 5 | 1993.412 | 57.41906 | 7.065916 | 0.25000000 | NA | 0.08333333 | DM | |
| 745 | 5 | 1993.496 | 57.50240 | 7.149250 | 0.33333333 | NA | 0.08333333 | DM | |
| 746 | 5 | 1993.579 | 57.58573 | 7.232583 | 0.41666667 | NA | 0.08333333 | DM | |
| 747 | 5 | 1993.662 | 57.66906 | 7.315916 | 0.50000000 | NA | 0.08333333 | DM | |
| 748 | 5 | 1993.746 | 57.75240 | 7.399250 | 0.58333333 | NA | 0.08333333 | DM | |
| 749 | 5 | 1993.829 | 57.83573 | 7.482583 | 0.66666667 | NA | 0.08333333 | DM | |
| 750 | 5 | 1993.912 | 57.91906 | 7.565916 | 0.75000000 | NA | 0.04671458 | DM | 1s |
| 751 | 5 | 1993.959 | 57.96578 | 7.612631 | 0.79671458 | 0.00000000 | 0.03661875 | 1st CVD | 1s |
| 752 | 5 | 1993.996 | 58.00240 | 7.649250 | 0.83333333 | 0.03661875 | 0.08333333 | 1st CVD | 1s |

# An added time scale II

```
753      5 1994.079 58.08573  7.732583 0.91666667 0.11995209 0.08333333 1st CVD 1s
754      5 1994.162 58.16906  7.815916 1.00000000 0.20328542 0.08333333 1st CVD 1s
755      5 1994.246 58.25240  7.899250 1.08333333 0.28661875 0.08333333 1st CVD 1s
756      5 1994.329 58.33573  7.982583 1.16666667 0.36995209 0.08333333 1st CVD 1s
757      5 1994.412 58.41906  8.065916 1.25000000 0.45328542 0.08333333 1st CVD 1s
758      5 1994.496 58.50240  8.149250 1.33333333 0.53661875 0.08333333 1st CVD 1s
759      5 1994.579 58.58573  8.232583 1.41666667 0.61995209 0.07819986 1st CVD 2r
760      5 1994.657 58.66393  8.310783 1.49486653 0.69815195 0.00513347 2nd CVD 2r
761      5 1994.662 58.66906  8.315916 1.50000000 0.70328542 0.08333333 2nd CVD 2r
762      5 1994.746 58.75240  8.399250 1.58333333 0.78661875 0.08333333 2nd CVD 2r
763      5 1994.829 58.83573  8.482583 1.66666667 0.86995209 0.08333333 2nd CVD 2r
764      5 1994.912 58.91906  8.565916 1.75000000 0.95328542 0.08333333 2nd CVD 2r
765      5 1994.996 59.00240  8.649250 1.83333333 1.03661875 0.08333333 2nd CVD 2r
766      5 1995.079 59.08573  8.732583 1.91666667 1.11995209 0.08333333 2nd CVD 2r
767      5 1995.162 59.16906  8.815916 2.00000000 1.20328542 0.08333333 2nd CVD 2r
768      5 1995.246 59.25240  8.899250 2.08333333 1.28661875 0.08333333 2nd CVD 2r
769      5 1995.329 59.33573  8.982583 2.16666667 1.36995209 0.08333333 2nd CVD 2r
770      5 1995.412 59.41906  9.065916 2.25000000 1.45328542 0.08333333 2nd CVD 2r
771      5 1995.496 59.50240  9.149250 2.33333333 1.53661875 0.08333333 2nd CVD 2r
```



# An added time scale III

```
772      5 1995.579 59.58573  9.232583 2.41666667 1.61995209 0.08333333 2nd CVD 2r
773      5 1995.662 59.66906  9.315916 2.50000000 1.70328542 0.08333333 2nd CVD 2r
774      5 1995.746 59.75240  9.399250 2.58333333 1.78661875 0.08333333 2nd CVD 2r
775      5 1995.829 59.83573  9.482583 2.66666667 1.86995209 0.08333333 2nd CVD 2r
776      5 1995.912 59.91906  9.565916 2.75000000 1.95328542 0.08333333 2nd CVD 2r
777      5 1995.996 60.00240  9.649250 2.83333333 2.03661875 0.08333333 2nd CVD 2r
778      5 1996.079 60.08573  9.732583 2.91666667 2.11995209 0.08333333 2nd CVD 2r
779      5 1996.162 60.16906  9.815916 3.00000000 2.20328542 0.08333333 2nd CVD 2r
780      5 1996.246 60.25240  9.899250 3.08333333 2.28661875 0.08333333 2nd CVD 2r
781      5 1996.329 60.33573  9.982583 3.16666667 2.36995209 0.08333333 2nd CVD 2r
782      5 1996.412 60.41906 10.065916 3.25000000 2.45328542 0.08333333 2nd CVD 2r
783      5 1996.496 60.50240 10.149250 3.33333333 2.53661875 0.08333333 2nd CVD 2r
784      5 1996.579 60.58573 10.232583 3.41666667 2.61995209 0.08333333 2nd CVD 2r
785      5 1996.662 60.66906 10.315916 3.50000000 2.70328542 0.08333333 2nd CVD 2r
786      5 1996.746 60.75240 10.399250 3.58333333 2.78661875 0.08333333 2nd CVD 2r
787      5 1996.829 60.83573 10.482583 3.66666667 2.86995209 0.08333333 2nd CVD 2r
788      5 1996.912 60.91906 10.565916 3.75000000 2.95328542 0.08333333 2nd CVD 2r
789      5 1996.996 61.00240 10.649250 3.83333333 3.03661875 0.08333333 2nd CVD 2r
790      5 1997.079 61.08573 10.732583 3.91666667 3.11995209 0.08333333 2nd CVD 2r
```



# An added time scale IV

```
791      5 1997.162 61.16906 10.815916 4.00000000 3.20328542 0.08333333 2nd CVD 2r
792      5 1997.246 61.25240 10.899250 4.08333333 3.28661875 0.08333333 2nd CVD 2r
793      5 1997.329 61.33573 10.982583 4.16666667 3.36995209 0.08333333 2nd CVD 2r
794      5 1997.412 61.41906 11.065916 4.25000000 3.45328542 0.08333333 2nd CVD 2r
795      5 1997.496 61.50240 11.149250 4.33333333 3.53661875 0.08333333 2nd CVD 2r
796      5 1997.579 61.58573 11.232583 4.41666667 3.61995209 0.08333333 2nd CVD 2r
797      5 1997.662 61.66906 11.315916 4.50000000 3.70328542 0.08333333 2nd CVD 2r
798      5 1997.746 61.75240 11.399250 4.58333333 3.78661875 0.08333333 2nd CVD 2r
799      5 1997.829 61.83573 11.482583 4.66666667 3.86995209 0.08333333 2nd CVD 2r
800      5 1997.912 61.91906 11.565916 4.75000000 3.95328542 0.08333333 2nd CVD 2r
801      5 1997.996 62.00240 11.649250 4.83333333 4.03661875 0.05099247 2nd CVD  I
```

# Representation of multistate FU: Lexis I

```
lex.id      per      age       dur      tsb    tfCVD   lex.dur lex.Cst lex.Xst
     5 1993.162 57.16906  6.815916 0.000000       NA 0.0833333      DM      DM
     5 1993.246 57.25240  6.899250 0.083333       NA 0.0833333      DM      DM
...
     5 1993.912 57.91906  7.565916 0.750000       NA 0.0467145      DM 1st CVD
     5 1993.959 57.96578  7.612631 0.796714 0.000000 0.0366187 1st CVD 1st CVD
     5 1993.996 58.00240  7.649250 0.833333 0.036618 0.0833333 1st CVD 1st CVD
...
     5 1994.579 58.58573  8.232583 1.416666 0.619952 0.0781998 1st CVD 2nd CVD
     5 1994.657 58.66393  8.310783 1.494866 0.698151 0.0051334 2nd CVD 2nd CVD
...
     5 1997.912 61.91906 11.565916 4.750000 3.953285 0.0833333 2nd CVD 2nd CVD
     5 1997.996 62.00240 11.649250 4.833333 4.036618 0.0509924 2nd CVD  Dth(2)
```



# Modeling mortality rates in Lexis objects

```
> dlev <- c("D(no CVD)", "D(1 CVD)",  "D(2 CVD)",  "D(3+ CVD)")
> #
> m0 <- glm( (lex.Xst %in% dlev ) ~
+            Ns( tsb, knots=d.kn ) + lex.Cst + allocation,
+            offset = log(lex.dur),
+            family = poisson,
+            data = S1 )
> #
> m1  <- update( m0, . ~ . + sex + age )                        # the real model
> #
> m1i <- update( m1, . ~ . - allocation + allocation:lex.Cst )
> #
> # Test interaction
> anova( m1i, m1, test="Chisq" )
```



# Modeling CVD rates in Lexis objects

```
> clev <- c("1st CVD","2nd CVD","3+ CVD")
> #
> c0 <- glm( ( (lex.Xst %in% clev) & (lex.Cst!=lex.Xst) ) ~
+            Ns( tsb, knots=d.kn ) + lex.Cst + allocation,
+            offset = log(lex.dur),
+            family = poisson,
+            data = subset( S1, lex.Cst!="3+ CVD" ) )
> #
> c1  <- update( c0, . ~ . + sex + age )
> #
> c1i <- update( c1, . ~ . - allocation + allocation:lex.Cst )
> #
> c1p <- update( c1, . ~ . + allocation:tsb )
> #
> # Test interaction & PH
> anova( c1i, c1, c1p, test="Chisq" )
```

# Simulation: transition object

```
> TM1 <- list( "DM" = list( "D(no CVD)" = m1,
+                           "1st CVD" = c1 ),
+         "1st CVD" = list(  "D(1 CVD)" = m1,
+                           "2nd CVD" = c1 ),
+         "2nd CVD" = list(  "D(2 CVD)" = m1,
+                            "3+ CVD" = c1 ),
+          "3+ CVD" = list( "D(3+ CVD)" = m1 ) )
```



# Simulation: baseline (init) object

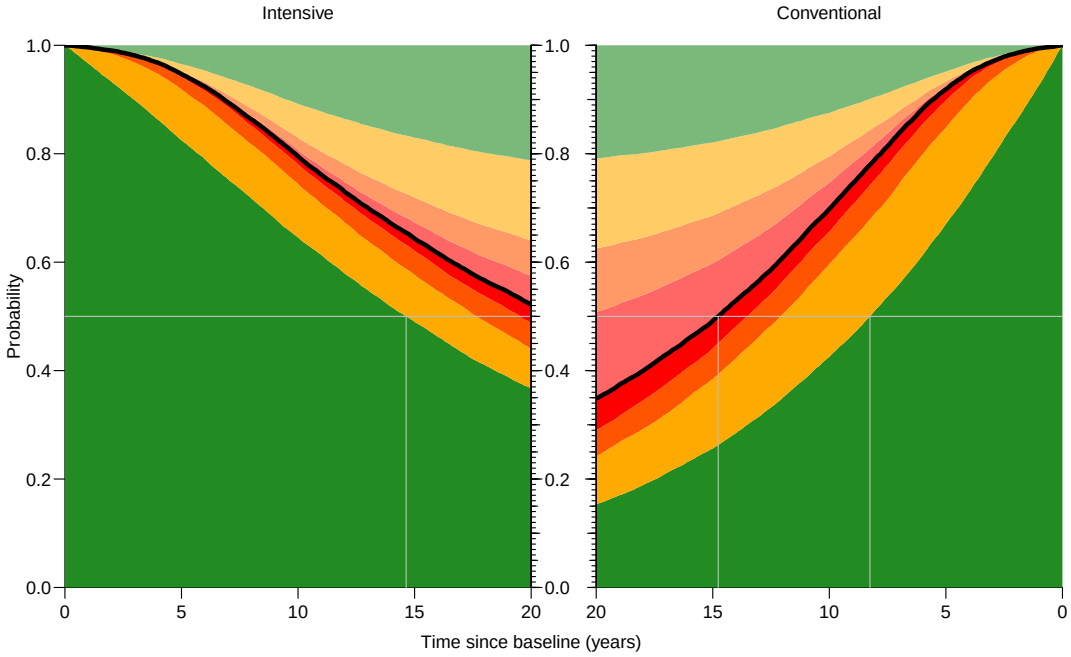
```
> St2 <- subset( L1, select = c(timeScales(L1),"lex.Cst","lex.dur",
+                               "allocation","sex") )
> levels(St2$lex.Cst) <- levels(S1$lex.Cst)
> nr <- nrow(St2)
> St2[1:nr,"tsb"] <- 0
> St2[1:nr,"lex.Cst"] <- levels(St2$lex.Cst)[1]
> St2[1:nr,"allocation"] <- levels(S1$allocation)[1]
> St2 <- rbind( St2, St2 )
> St2[1:nr,"allocation"] <- levels(S1$allocation)[2]
```

St2 is a copy of the original date with `allocation="Intensive"` plus a copy of the original date with `allocation="Conventional"`

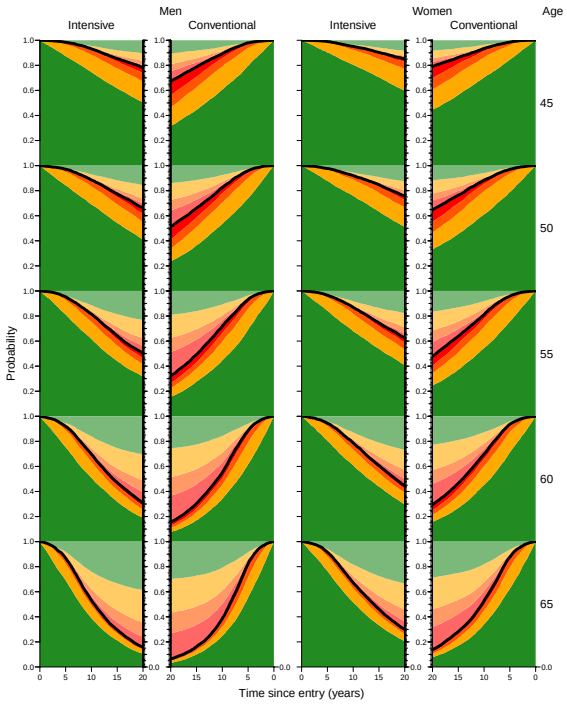


# Simylating a follow up dataset

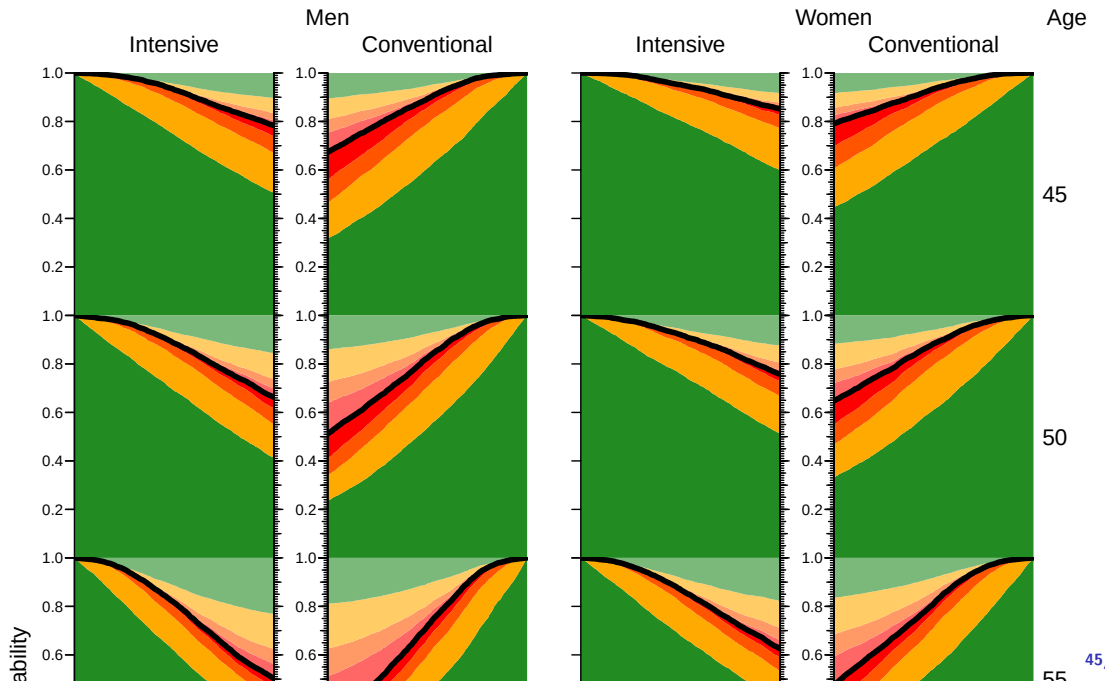
```
> sim2 <- simLexis( Tr = TM1,
+                   init = St2,
+                      N = 100,
+              time.pts = seq(0,21,0.1) )
> StC <- nState( subset(sim2,allocation=="Conventional"),
+                at=seq(0,20,0.1), from=0, time.scale="tsb" )
> StI <- nState( subset(sim2,allocation=="Intensive"),
+                at=seq(0,20,0.1), from=0, time.scale="tsb" )
> prm <- c(1:4,8:5)
> pC <- pState( StC, perm=prm )
> pI <- pState( StI, perm=prm )
```


Intensive
Conventional
Probability
Time since baseline (years)




Men
Women
Age
Intensive
Conventional
Probability
Time since entry (years)

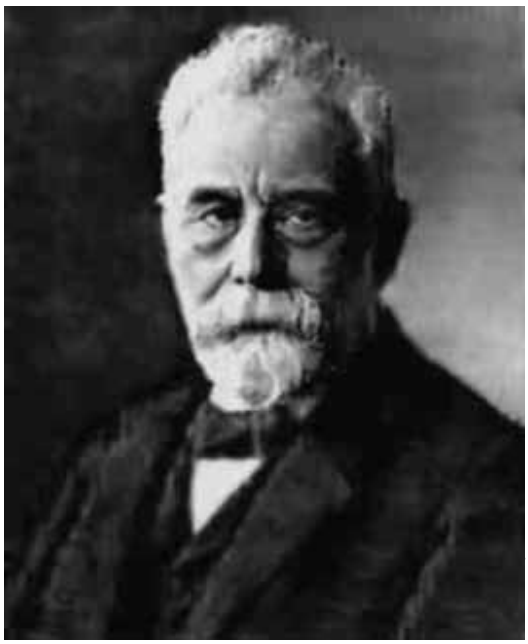


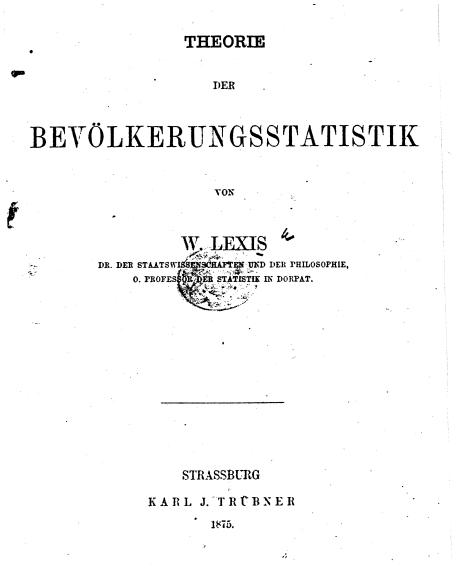

Men
Women
Age
Intensive
Conventional

## History

- `Epi` package grew out of "Statistical Practice in Epidemiology with R", annually since 2002 in Tartu Estonia
- `Lexis` machinery conceived by Martyn Plummer, IARC
- Naming originally by David Clayton & Michael Hills, `stlexis` in Stata, later renamed `stsplit`
- David Claytion wrote a `lexis` function for the `Epi` package. Obsolete now.






THEORIE

DER

BEVÖLKERUNGSSTATISTIK

VON

W. LEXIS

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

STRASSBURG

KARL J. TRÜBNER

1875.



## Summary

- Proper representation of multistate data essential: States, transitions, risk time
- Readable modeling code
- Calculation of state probabilities requires a simulation in any realistic situation
- Examples of use in: `http://bendixcarstensen.com/AdvCoh/Lexis-ex/`

# Thanks for your attention