Mortality among Type 2 patients at Steno Diabetes Center

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

1 Data preparation .................................................. 1
  1.1 Introduction ................................................. 1
  1.2 Data preparation ........................................... 1
    1.2.1 Reading with R .................................. 3
    1.2.2 Causes of death .................................. 8
  1.3 Lexis object for analysis of overall mortality rates .......... 9
    1.3.1 Raw mortality by calendar year ................. 9
  1.4 Lexis object for analysis of cause-specific mortality ...... 13
  1.5 Base tables ................................................. 16

2 Mortality by cause of death .................................. 18
  2.1 Statistical analysis ....................................... 21
    2.1.1 Setup ........................................... 21
  2.2 Age and date of follow-up ................................ 22
  2.3 Including diabetes duration ................................ 25

3 All-cause mortality by complication status .................. 29
  3.1 Setup ...................................................... 29
  3.2 Duration of complications ................................ 33
  3.3 Splitting follow-up for Poisson analysis .................. 34
  3.4 Analysis of duration variables ............................ 35
    3.4.1 Spline setup .................................... 36
  3.5 Including population mortality rates ....................... 39
  3.6 Modeling all cause mortality rates ........................ 40
    3.6.1 Timescale selection ............................... 41
  3.7 Estimation and extraction of the results .................... 42
    3.7.1 Graphs and tables for the paper ................. 54
    3.7.2 Results from the naive model .................... 59
Chapter 1

Data preparation

1.1 Introduction

This report concerns the mortality among type 2 patients only at Steno Diabetes Center in the period where the computerized patients records (EPJ) are available, that is the period 1.9.2001–15.9.2010.

The records from EPJ were linked to the Central Person Register CPR and the cause of death register, so we know the date of death for those who died.

For the years 2009 ff. the cause of death is not known, owing to a backlog of death certificates, but the date of death is known up till 15th September 2010. This report is concerned only with the overall mortality rates, and thus comprises follow-up from 1.1.2002–15.9.2010.

1.2 Data preparation

We initially read the data from a SAS-file, and exported it to the file ./data/mcompl.xpt:

```
1 "Program: getit.sas" 10:49 Tuesday, January 29, 2013

NOTE: Copyright (c) 2002-2008 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) Proprietary Software 9.2 (TS2M3)
Licensed to NOVO NORDISK - BASIC PACKAGE, Site 50800704.
NOTE: This session is executing on the W32_VSPRO platform.

NOTE: SAS initialization used:
real time 2.58 seconds
cpu time 0.37 seconds

NOTE: AUTOEXEC processing beginning; file is c:\stat\sas\autoexec.sas.

-------------------------------------------------------------------------------
C:\Bendix\Steno\MaEJ\EPJ-dod\sas\getit.sas
-------------------------------------------------------------------------------

NOTE: Libref HER was successfully assigned as follows:
Engine: V9
Physical Name: C:\Bendix\Steno\MaEJ\EPJ-dod\sas

NOTE: Libref DATA was successfully assigned as follows:
Engine: V9
Physical Name: p:\MAEJ\SAS data\SDC mortality

NOTE: AUTOEXEC processing completed.

1 options noplerr ;
2 libname maej "p:\MAEJ\SAS data\SDC mortality" ;
3 NOTE: Libref MAEJ was successfully assigned as follows:
Engine: V9
Physical Name: p:\MAEJ\SAS data\SDC mortality
```
title1 "Base dataset - merge of EPJ, NPR and CoDR";
proc contents data=maej.compl;
run;
NOTE: PROCEDURE CONTENTS used (Total process time):
real time 0.07 seconds
cpu time 0.06 seconds
NOTE: The PROCEDURE CONTENTS printed page 1.

* Identify fishy records;
data oops nodm notp late compl;
set maej.compl;
if doBth gt doDM gt .z then do;
   put "This was changed from: " doBth= ddmmyy10.
doDM= doBth + 90;
   put " to: " doBth= ddmmyy10.
doDM= ddmmyy10.
end;
if doBth lt "01JAN1900"d then do;
   put "This was changed from: " doBth= ddmmyy10.
doDM= ddmmyy10.
doBth = mdy( month(doBth), day(doBth), year(doBth)+100 );
   put " to: " doBth= ddmmyy10.
doDM= ddmmyy10.
end;
if ( nmiss(doCVD,CVD) eq 1 or
   nmiss(doDR ,DR ) eq 1 or
   nmiss(doNef,Nef) eq 1 or
   nmiss(doNeu,Neu) eq 1 or
   ( doDth gt .z and CoDth eq "" ) or
   ( doDth le .z and CoDth ne "" ) or
   nmiss(sex,doBth,entry,exit) gt 0 or
   doBth gt entry or
   entry gt exit ) then output oops;
else if doDM le .z then output nodm;
else if DMtype eq "" then output notp;
else if entry gt "31DEC2010"d then output late;
else output compl;
run;

This was changed from: doBTH=06/07/1970 doDM=01/01/1970
to: doBTH=06/07/1970 doDM=04/10/1970
This was changed from: doBTH=22/05/1964 doDM=01/01/1964
to: doBTH=22/05/1964 doDM=20/08/1964
This was changed from: doBTH=25/12/1977 doDM=01/01/1977
to: doBTH=25/12/1977 doDM=25/03/1978
This was changed from: doBTH=29/01/1887 doDM=01/01/1993
to: doBTH=29/01/1887 doDM=01/01/1993
NOTE: There were 11424 observations read from the data set MAEJ.COMPL.
NOTE: The data set WORK.OOPS has 2 observations and 16 variables.
NOTE: The data set WORK.NODM has 542 observations and 16 variables.
NOTE: The data set WORK.NOTP has 12 observations and 16 variables.
NOTE: The data set WORK.LATE has 137 observations and 16 variables.
NOTE: The data set WORK.COMPL has 10731 observations and 16 variables.
NOTE: DATA statement used (Total process time):
real time 0.20 seconds
cpu time 0.04 seconds

*proc print data=maej.compl;
var sex dmtype dobth dodm exit entry doCVD CVD doDR DR doNef Nef doNeu Neu doDth CoDth;
format dobth dodm exit entry doCVD CVD doDR DR doNef Nef doNeu Neu doDth CoDth rmmyy8. ;
run;

NOTE: There were 2 observations read from the data set WORK.OOPS.
NOTE: PROCEDURE PRINT used (Total process time):
real time 0.00 seconds
cpu time 0.01 seconds

* /*
title1 "nodm";
proc print data=maej.compl;
var sex dmtype dobth dodm exit entry doCVD CVD doDR DR doNef Nef doNeu Neu doDth CoDth;
format dobth dodm exit entry doCVD CVD doDR DR doNef Nef doNeu Neu doDth CoDth rmmyy8. ;
run;

NOTE: There were 2 observations read from the data set WORK.OOPS.
NOTE: PROCEDURE PRINT used (Total process time):
real time 0.00 seconds
cpu time 0.01 seconds

* /*
title1 "notp";
proc print data=maej.compl;
var sex dmtype dobth dodm exit entry doCVD CVD doDR DR doNef Nef doNeu Neu doDth CoDth;
format dobth dodm exit entry doCVD CVD doDR DR doNef Nef doNeu Neu doDth CoDth rmmyy8. ;
run;
194
Data preparation

1.2 Data preparation

```sas
/*
libname xptout xport './data/compl.xpt';
NOTE: Libref XPTOUT was successfully assigned as follows:
Engine: XPORT
Physical Name: C:\Bendix\Steno\MaEJ\EPJ-dod\data\compl.xpt
proc copy in=work out=xptout memtype=data;
select compl;
run;
NOTE: Copying WORK.COMPL to XPTOUT.COMPL (memtype=DATA).
NOTE: There were 10731 observations read from the data set WORK.COMPL.
NOTE: The data set XPTOUT.COMPL has 10731 observations and 16 variables.
NOTE: PROCEDURE COPY used (Total process time):
   real time 0.22 seconds
cpu time 0.04 seconds

NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414
NOTE: The SAS System used:
   real time 3.45 seconds
cpu time 0.60 seconds
```

The CONTENTS Procedure

```
Data Set Name MAEJ.COMPL Observations 11424
Member Type DATA Variables 16
Engine V9 Indexes 0
Created 28. januar 2013 mandag 20:54:14 Observation Length 128
Last Modified 28. januar 2013 mandag 20:54:14 Deleted Observations 0
Protection Compressed NO
Data Set Type Sorted NO
Label
Data Representation WINDOWS_32
Encoding vlatini Western (Windows)
```

1.2.1 Reading with R

We first load the package needed to read the data:
R version 3.0.1 (2013-05-16)
Platform: i386-w64-mingw32/i386 (32-bit)

attached base packages:
[1] splines  utils  datasets  graphics  grDevices  stats  methods  base

other attached packages:
[1] Epi_1.1.57 foreign_0.8-53

loaded via a namespace (and not attached):
[1] tools_3.0.1

Then we read the data from the SAS export file, and restrict to T2 patients:

```r
ten <- read.xport( "./data/compl.xpt" )
ten <- subset( ten, dmtype=="T2", select=-2 )
```

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>T2</td>
</tr>
<tr>
<td>4855</td>
<td>5876</td>
</tr>
</tbody>
</table>

Then we define sex as a factor and transform a format of fractional years. Since date of diagnosis is only given as a year we pick the date of diagnosis randomly in the year, but so that it is not after entry and at least a month before death:
Data preparation

```r
> epj$sex <- factor(epj$sex, labels=c("M","F") )
> dnum <- c( grep("entry", names(epj) ),
+ grep("exit", names(epj) ),
+ grep("do", names(epj) ) )
> names( epj )[dnum]

[1] "entry" "exit" "dodth" "dobth" "dodm" "docvd" "dodr" "donef" "doneu"
```

```r
> for( i in dnum ) epj[,i] <- epj[,i]/365.25 + 1960
> set.seed( 783459876 )
> epj$dodm <- pmin( epj$dodm + runif( nrow(epj) ),
+ epj$entry,
+ epj$dodth-1/12,
+ na.rm=TRUE )
> options(digits=6)
> head( epj )
```

```r
dodth sex dobth dodm exit entry codth docvd cvd dodr dr donef nef doneu neu
5 NA M 1930 1987.76 2002.43 2011.61 1 NA NA 1993.75 1 1993.75 1
```

```r
> options(digits=8)
> with( epj, ftable( addmargins( table( sex,
+ Death=floor(dodth),
+ Entry=floor(entry),
+ useNA="ifany" ) ) )
```

Once we read data, we can get a quick overview of the dataset check that entry and death dates are in the correct relation to each other:

```r
> with( epj, ftable( addmargins( table( sex,
+ Death=floor(dodth),
+ Entry=floor(entry),
+ useNA="ifany" ) ) )
```

<table>
<thead>
<tr>
<th>Entry</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>sex</td>
<td>M</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2001</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>2002</td>
<td>71</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>2003</td>
<td>72</td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>78</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td>70</td>
<td>12</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td>59</td>
<td>11</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>2007</td>
<td>42</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>54</td>
<td>10</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>65</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>59</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>7</td>
<td>8</td>
<td>2</td>
<td>0</td>
<td>96</td>
</tr>
<tr>
<td>Sum</td>
<td></td>
<td>1527</td>
<td>222</td>
<td>231</td>
<td>268</td>
<td>218</td>
<td>188</td>
<td>203</td>
<td>177</td>
<td>204</td>
<td>3441</td>
</tr>
<tr>
<td>sex</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2001</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2002</td>
<td>46</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>2003</td>
<td>38</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>55</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td>45</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td>49</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>2007</td>
<td>44</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>46</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>32</td>
<td>7</td>
<td>1</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>35</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>57</td>
</tr>
<tr>
<td>Sum</td>
<td></td>
<td>1527</td>
<td>222</td>
<td>231</td>
<td>268</td>
<td>218</td>
<td>188</td>
<td>203</td>
<td>177</td>
<td>204</td>
<td>3441</td>
</tr>
<tr>
<td>sex</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>939</td>
<td>157</td>
<td>177</td>
<td>240</td>
<td>189</td>
<td>167</td>
<td>182</td>
<td>165</td>
<td>203</td>
<td>202</td>
<td>2621</td>
</tr>
<tr>
<td>Sum</td>
<td>1527</td>
<td>222</td>
<td>231</td>
<td>268</td>
<td>218</td>
<td>188</td>
<td>203</td>
<td>177</td>
<td>204</td>
<td>3441</td>
<td></td>
</tr>
</tbody>
</table>
```
To see how the follow-up is working we just check how dates of diagnosis resp. complications and date of death are distributed, and also how date of birth and date of diagnosis look, as well as how date of diagnosis and date of exit look:

```r
> par( mfrow=c(3,2), mar=c(3,3,2,1), mgp=c(3,1,0)/1.6, bty="n" )
> with( epj, +   hist(dodth,breaks=seq(2001,2012,1/12),col="gray",main="All cause"))
> abline(v=2001:2012,col="red")
> with( subset( epj, codth != "" ), +   hist(dodth,breaks=seq(2001,2012,1/12),col="gray",main="Cause known" ) )
> abline(v=2001:2012,col="red")
> with( epj, plot( dobth, dodm, pch=16, cex=0.8 ) )
> abline(0,1,col="red")
> with( epj, plot( exit, dodm, pch=16, cex=0.8 ) )
> abline(0,1,col="red")
> with( epj, plot( dodm, donef, pch=16, cex=0.8 ) )
> abline(0,1,col="red")
> with( epj, plot( exit, donef, pch=16, cex=0.8 ) )
> abline(0,1,col="red")
```

From figure 1.1 it is clear that the exit date for both all cause mortality analysis and for the cause-specific analyses should be 2010-12-31; the maximal date of death in the data frame is 2010-12-31 So we just check if all persons enter and exits correctly:

```r
> all.exit <- cal.yr( "2011-01-01" )
> with( epj, ftable( addmargins( +   table( sex, +   deathOK = dodth < all.exit, +   entryOK = entry < all.exit, +   useNA="ifany" ), margin=1:2 ), +   col.vars=c(1,3) ) )
```

```
   sex M F Sum
entryOK TRUE TRUE TRUE
deathOK TRUE 820 553 1373
NA 2621 1882 4503
Sum 3441 2435 5876
```

From the two bottom panels of figure 1.1 we discover an anomaly in the dates of nephropathy:

```r
> tt <- with( epj, addmargins( table(DN=round(donef,3), +   Dead=!is.na(dodth), +   exclude=NULL ) )
> tt[tt[,"Sum"]>5,]
```
Figure 1.1: Histogram of dates of death for all known deaths and for deaths where a cause is known. The bottom 4 panels have plots of the date of diagnosis resp. complications versus date of birth and date of exit, with a red line indicating the identity (so all points should preferably be on the same side of this).
Thus it seems that some sort of update of the patients’ nephropathy status has taken place in the fall of 1993 and 2005, and maybe even that the update has been restricted to patients alive at some later state. This means that the nephropathy status presumably is recorded with different precision over the period, such that those with dates recorded at these two dates are patients that are included with DN because of some status, and hence may have their dates of DN recorded earlier in the course of the complications history than the other patients with DN. This would only have the effect that DN is recorded with uncertainty. In future studies it would be prudent to define nephropathy status directly from the clinical recordings.

### 1.2.2 Causes of death

Finally we define CoD (Cause of Death) as a factor with 4 causes of death and “Alive” for those not yet dead, and make a check that it all went well:

```r
> with( epj, table(codth) )

<table>
<thead>
<tr>
<th>codth</th>
<th>accidents</th>
<th>acute DM</th>
<th>Cancer</th>
<th>CVD</th>
<th>GI</th>
<th>Infection</th>
<th>kidney</th>
<th>Lung</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>4503</td>
<td>17</td>
<td>11</td>
<td>288</td>
<td>493</td>
<td>67</td>
<td>91</td>
<td>34</td>
<td>84</td>
<td>288</td>
</tr>
</tbody>
</table>

> # Capitalize the causes of death
> levels( epj$codth )[-1] <-
  + sapply( strsplit(levels(epj$codth),""),
  + function(x) { x[1] <- toupper(x[1]) ; paste(x,collapse="") } )[-1]
> epj$CoD <- Relevel( epj$codth, c(1,5,4,3,8,9,6,7,2) )
> epj$CoD <- Relevel( epj$codth, list("Alive"=1,2,3,"Other"=4:10) )
> with( epj, table( codth, CoD ) )

<table>
<thead>
<tr>
<th>CoD</th>
<th>Alive</th>
<th>CVD</th>
<th>Cancer</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>4503</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CVD</td>
<td>0</td>
<td>0</td>
<td>288</td>
<td>0</td>
</tr>
<tr>
<td>Cancer</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Acute DM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>34</td>
</tr>
<tr>
<td>Kidney</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>Lung</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>67</td>
</tr>
<tr>
<td>GI</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>91</td>
</tr>
<tr>
<td>Infection</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Accidents</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>288</td>
</tr>
</tbody>
</table>

> with( epj, ftable( addmargins( table( sex,
  + doDTH=floor(dodth), CoD,
  + useNA="ifany" ),
  + margin=2:3), row.vars=2 ) )
1.3 Lexis object for analysis of overall mortality rates

In order to analyze all cause mortality of diabetes patients in SDC we set up a Lexis object which holds the follow-up time on the timescales age, diabetes duration and calendar time. Hence we also exclude those without a date of diabetes diagnosis (and those who have a date of entry before date of diabetes):

```r
L1 <- Lexis( entry = list( age = entry-dobth, + Ddur = entry-dodm, + per = entry ), + exit = list( per = pmin(exit,all.exit,na.rm=TRUE) ), + exit.status = factor( !is.na(dodth), labels=c("Alive","Dead") ), + data = epj )
```

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

```r
> summary.Lexis( L1, by="sex" )
```

### $M$

**Transitions:**

**To**

**From** Alive Dead **Records:** Events **Risk time:** Persons:

Alive 2621 820 3441 820 20144.39 20144.39

### $F$

**Transitions:**

**To**

**From** Alive Dead **Records:** Events **Risk time:** Persons:

Alive 1882 553 2435 553 14597.89 14597.89

1.3.1 Raw mortality by calendar year

We now check how the empirical mortality rates look by calendar year after this grooming. To this end we split the follow-up in 3-month intervals by calendar time:

```r
S1 <- splitLexis( L1, time.scale="per", breaks=1995+seq(0,20,1/4) )
> summary( S1 )
```
Transitions:

<table>
<thead>
<tr>
<th>From</th>
<th>Alive</th>
<th>Dead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Records</td>
<td>142246</td>
<td>1373</td>
</tr>
<tr>
<td>Events</td>
<td>143619</td>
<td>1573</td>
</tr>
<tr>
<td>Risk</td>
<td>34742.28</td>
<td>5876</td>
</tr>
</tbody>
</table>

A quick tabulation reveals that early mortality rates are really low; tabulation by 3-month period of follow-up gives:

```r
> DY <- xtabs( cbind(D=lex.Xst!="Alive", + Y=lex.dur, + rate=lex.dur) + I(floor(per*4)/4) + sex, + data=S1 )
> DY[,,"rate"] <- DY[,,"D"]/DY[,,"Y"]*100
> round( ftable( DY, row.vars=1 ), 1 )
```

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>D Y rate</td>
<td>D Y rate</td>
<td></td>
</tr>
<tr>
<td>I(floor(per * 4)/4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>0.0 12.4</td>
<td>0.0 9.1</td>
</tr>
<tr>
<td>2001.25</td>
<td>3.0 330.5</td>
<td>0.0 312.2</td>
</tr>
<tr>
<td>2001.5</td>
<td>5.0 347.1</td>
<td>1.4 327.1</td>
</tr>
<tr>
<td>2001.75</td>
<td>10.0 367.6</td>
<td>2.7 347.6</td>
</tr>
<tr>
<td>2002</td>
<td>17.0 383.6</td>
<td>4.4 367.6</td>
</tr>
<tr>
<td>2002.25</td>
<td>22.0 400.4</td>
<td>5.5 383.6</td>
</tr>
<tr>
<td>2002.5</td>
<td>18.0 410.3</td>
<td>4.4 391.0</td>
</tr>
<tr>
<td>2003</td>
<td>17.0 420.2</td>
<td>4.0 391.0</td>
</tr>
<tr>
<td>2003.25</td>
<td>25.0 430.4</td>
<td>5.8 410.3</td>
</tr>
<tr>
<td>2003.5</td>
<td>19.0 437.5</td>
<td>4.3 410.3</td>
</tr>
<tr>
<td>2003.75</td>
<td>26.0 447.5</td>
<td>5.8 410.3</td>
</tr>
<tr>
<td>2004</td>
<td>27.0 457.3</td>
<td>5.9 410.3</td>
</tr>
<tr>
<td>2004.25</td>
<td>28.0 468.5</td>
<td>6.0 420.2</td>
</tr>
<tr>
<td>2004.5</td>
<td>14.0 477.5</td>
<td>2.9 410.3</td>
</tr>
<tr>
<td>2004.75</td>
<td>20.0 489.7</td>
<td>4.1 410.3</td>
</tr>
<tr>
<td>2005</td>
<td>23.0 501.0</td>
<td>4.6 410.3</td>
</tr>
<tr>
<td>2005.25</td>
<td>23.0 508.7</td>
<td>4.5 410.3</td>
</tr>
<tr>
<td>2005.5</td>
<td>24.0 515.2</td>
<td>4.7 410.3</td>
</tr>
<tr>
<td>2005.75</td>
<td>25.0 522.7</td>
<td>4.8 410.3</td>
</tr>
<tr>
<td>2006</td>
<td>25.0 531.4</td>
<td>4.7 410.3</td>
</tr>
<tr>
<td>2006.25</td>
<td>21.0 537.1</td>
<td>3.9 410.3</td>
</tr>
<tr>
<td>2006.5</td>
<td>19.0 542.9</td>
<td>3.5 410.3</td>
</tr>
<tr>
<td>2006.75</td>
<td>28.0 546.0</td>
<td>5.1 410.3</td>
</tr>
<tr>
<td>2007</td>
<td>26.0 553.6</td>
<td>4.7 410.3</td>
</tr>
<tr>
<td>2007.25</td>
<td>13.0 562.2</td>
<td>2.3 410.3</td>
</tr>
<tr>
<td>2007.5</td>
<td>19.0 569.3</td>
<td>3.3 410.3</td>
</tr>
<tr>
<td>2007.75</td>
<td>19.0 577.8</td>
<td>3.3 410.3</td>
</tr>
<tr>
<td>2008</td>
<td>21.0 585.5</td>
<td>3.6 410.3</td>
</tr>
<tr>
<td>2008.25</td>
<td>21.0 590.4</td>
<td>3.6 410.3</td>
</tr>
<tr>
<td>2008.5</td>
<td>18.0 593.9</td>
<td>3.0 410.3</td>
</tr>
<tr>
<td>2008.75</td>
<td>31.0 599.7</td>
<td>5.2 410.3</td>
</tr>
<tr>
<td>2009</td>
<td>27.0 606.8</td>
<td>4.4 410.3</td>
</tr>
<tr>
<td>2009.25</td>
<td>29.0 613.8</td>
<td>4.7 410.3</td>
</tr>
<tr>
<td>2009.5</td>
<td>22.0 618.0</td>
<td>3.6 410.3</td>
</tr>
<tr>
<td>2009.75</td>
<td>24.0 625.0</td>
<td>3.8 410.3</td>
</tr>
<tr>
<td>2010</td>
<td>23.0 633.1</td>
<td>3.6 410.3</td>
</tr>
<tr>
<td>2010.25</td>
<td>29.0 639.3</td>
<td>4.5 410.3</td>
</tr>
<tr>
<td>2010.5</td>
<td>23.0 644.4</td>
<td>3.6 410.3</td>
</tr>
<tr>
<td>2010.75</td>
<td>21.0 649.4</td>
<td>3.2 410.3</td>
</tr>
</tbody>
</table>
```

```r
> matplot( as.numeric(dimnames(DY)[[1]]), log="y", las=1, + xlab="Date", ylab="Raw mortality (% / year)", + DY[,,"rate"], type="l", lty=1, lwd=3, col=c("blue","red") )
> abline( v=seq(1998,2015,1/4), col=gray(0.9) )
> abline( v=seq(1998,2015,1) , col=gray(0.8) )
> box()
```
Figure 1.2: Raw mortality rates for T2 patients by 3-month periods; there is something missing prior to 2002.
A quick look at the tables or figure 1.2 shows that data seem incomplete prior to 2002, so to make sure that data are valid, we define entry to be at the start of 2002, and redefine the Lexis object:

```r
> epj$entry <- pmax( epj$entry, 2002 )
> L1 <- Lexis( entry = list( age = entry-dobth,
+ Ddur = entry-dodm,
+ per = entry ),
+ exit = list( per = pmin(exit,all.exit,na.rm=TRUE) ),
+ exit.status = factor( !is.na(dodth), labels=c("Alive","Dead") ),
+ data = subset( epj, entry < pmin(exit,all.exit,na.rm=TRUE) ) )
```

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

```r
> summary( L1, by="sex" )
```

$M

Transitions:
To
From Alive Dead Records: Events: Risk time: Persons:
Alive 2621 802 3423 802 19086.81 3423

$F

Transitions:
To
From Alive Dead Records: Events: Risk time: Persons:
Alive 1882 539 2421 539 13826.64 2421

```r
> summary.data.frame( L1 )
```

<table>
<thead>
<tr>
<th>age</th>
<th>Ddur</th>
<th>per</th>
<th>lex.dur</th>
<th>lex.Cst</th>
<th>lex.Xst</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. :14.290</td>
<td>Min. :0.0000</td>
<td>Min. :2002.0</td>
<td>Min. :0.0068446</td>
<td>Alive :5844</td>
<td>Alive :4503</td>
</tr>
<tr>
<td>1st Qu. :52.389</td>
<td>1st Qu. :2.3210</td>
<td>1st Qu. :2002.0</td>
<td>1st Qu. :2.9226557</td>
<td>Dead : 0</td>
<td>Dead :1341</td>
</tr>
<tr>
<td>Mean :60.392</td>
<td>Mean :8.6015</td>
<td>Mean :2004.4</td>
<td>Mean :5.6320081</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd Qu. :69.454</td>
<td>3rd Qu. :12.8486</td>
<td>3rd Qu. :2006.5</td>
<td>3rd Qu. :8.9993155</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max. :96.523</td>
<td>Max. :51.4427</td>
<td>Max. :2011.0</td>
<td>Max. :8.9993155</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>lex.id</th>
<th>dodth</th>
<th>sex</th>
<th>dobth</th>
<th>dodm</th>
<th>exit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. :1.0</td>
<td>Min. :2002.0</td>
<td>Min. :1905.6</td>
<td>Min. :1950.6</td>
<td>Min. :2002.0</td>
<td></td>
</tr>
<tr>
<td>1st Qu. :1461.8</td>
<td>1st Qu. :2004.5</td>
<td>F:2421</td>
<td>1st Qu. :1934.4</td>
<td>1st Qu. :1990.5</td>
<td>1st Qu. :2012.4</td>
</tr>
<tr>
<td>Median :2922.5</td>
<td>Median :2006.7</td>
<td>Median :1943.6</td>
<td>Median :1997.3</td>
<td>Median :2012.4</td>
<td></td>
</tr>
<tr>
<td>Mean :2922.5</td>
<td>Mean :2006.7</td>
<td>Mean :1944.0</td>
<td>Mean :1995.8</td>
<td>Mean :2011.1</td>
<td></td>
</tr>
<tr>
<td>3rd Qu. :4383.2</td>
<td>3rd Qu. :2008.9</td>
<td>3rd Qu. :1952.3</td>
<td>3rd Qu. :2002.3</td>
<td>3rd Qu. :2012.4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>entry</th>
<th>codth</th>
<th>docvd</th>
<th>cvd</th>
<th>cvo</th>
<th>dr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Qu. :2002.0</td>
<td>CVD :473</td>
<td>1st Qu. :1996.8</td>
<td>1st Qu. :1</td>
<td>1st Qu. :1996.6</td>
<td>1st Qu. :1</td>
</tr>
<tr>
<td>Mean :2004.4</td>
<td>Cancer :285</td>
<td>Mean :2001.3</td>
<td>Mean :1</td>
<td>Mean :2001.5</td>
<td>Mean :1</td>
</tr>
<tr>
<td>3rd Qu. :2006.5</td>
<td>Infection :89</td>
<td>3rd Qu. :2005.1</td>
<td>3rd Qu. :1</td>
<td>3rd Qu. :2005.7</td>
<td>3rd Qu. :1</td>
</tr>
<tr>
<td>(Other) :126</td>
<td>NA's :1045</td>
<td>NA's :1045</td>
<td>NA's :2879</td>
<td>NA's :2879</td>
<td></td>
</tr>
<tr>
<td>donef</td>
<td>nef</td>
<td>doneu</td>
<td>neu</td>
<td>CoD</td>
<td></td>
</tr>
<tr>
<td>Min. :1992.2</td>
<td>Min. :1</td>
<td>Min. :1991.5</td>
<td>Min. :1</td>
<td>Alive :4503</td>
<td></td>
</tr>
<tr>
<td>1st Qu. :1997.2</td>
<td>1st Qu. :1</td>
<td>1st Qu. :1996.4</td>
<td>1st Qu. :1</td>
<td>CVD :473</td>
<td></td>
</tr>
<tr>
<td>Mean :2001.8</td>
<td>Mean :1</td>
<td>Mean :2000.8</td>
<td>Mean :1</td>
<td>Other :583</td>
<td></td>
</tr>
<tr>
<td>3rd Qu. :2005.9</td>
<td>3rd Qu. :1</td>
<td>3rd Qu. :2004.6</td>
<td>3rd Qu. :1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max. :2012.4</td>
<td>Max. :1</td>
<td>Max. :2012.4</td>
<td>Max. :1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA's :3686</td>
<td>NA's :3686</td>
<td>NA's :2950</td>
<td>NA's :2950</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We can now make an overview of the age and DM-duration distribution at the entry of the study:

```
> par( mfcol=c(2,2), mar=c(2,2,1,1), mgp=c(3,1,0)/1.6, las=1, oma=c(0,0,2,0) )
> yl <- c(0,190)
> with( subset( L1, sex=="M" ),
+     hist( age, breaks=0:101, col="blue", border="blue",
+           main="", ylim=yl, xlab="", ylab="" ) )
> with( subset(L1, sex=="F"),
+     hist( age, breaks=0:100, col="red", border="red",
+           main="", ylim=yl, xlab="", ylab="" ) )
> zz <-
+     with( subset( L1, sex=="M" ),
+           hist( Ddur, breaks=0:101, col="blue", border="blue",
+                 main="", ylim=yl, xlab="", ylab="" ) )
> text( 1.5, yl[2], zz$counts[1], font=2, adj=c(-0.1,0), col="blue" )
> zz <-
+     with( subset(L1, sex=="F"),
+           hist( Ddur, breaks=0:100, col="red", border="red",
+                 main="", ylim=yl, xlab="", ylab="" ) )
> text( 1.5, yl[2], zz$counts[1], font=2, adj=c(-0.1,0), col="red" )
> mtext( c("Age at entry","DM duration at entry"), side=3, line=0,
+        at=c(1,3)/4, outer=TRUE )
```

1.4 Lexis object for analysis of cause-specific mortality

This is completely parallel to the set-up above, except that we use the factor CoD as the exit variable.

```
> with( epj, table( floor(exit), CoD ) )

   CoD  Alive CVD Cancer Other
 2001     0   20    3    9
 2002     0   48   18   56
 2003     0   57   29   43
 2004     0   53   32   69
 2005     0   66   33   55
 2006     0   65   32   74
 2007     0   34   32   67
 2008     0   56   31   71
 2009     0   53   40   74
 2010     0   41   38   74
2012 4503     0     0     0

> with( epj, table( deathOK = dodth < all.exit,
+ entryOK = entry < all.exit, useNA="ifany" ) )

  entryOK
deathOK TRUE
    TRUE 1373
<NA> 4503
Figure 1.3: Entry age and duration for T2 patients at Steno.
Data preparation

1.4 Lexis object for analysis of cause-specific mortality

```r
> C1 <- Lexis( entry = list( age = entry-dobth, 
+ Ddur = entry-dodm, 
+ per = entry ), 
+ exit = list( per = pmin(exit,all.exit,na.rm=TRUE) ), 
+ exit.status = CoD, 
+ data = subset( epj, entry < pmin(exit,all.exit,na.rm=TRUE) ) )

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

> nrow( C1 )

[1] 5844

> summary( C1 )

Transitions:
To
From Alive CVD Cancer Other Records: Events: Risk time: Persons:
Alive 4503 473 285 583 5844 1341 32913.46 5844

> with( C1, ftable( addmargins( table( codth, sex, CoD ), 
+ margin=c(1,3) ), 
+ col.vars=2:3 ) )

<table>
<thead>
<tr>
<th>sex</th>
<th>CoD Alive</th>
<th>CVD</th>
<th>Cancer</th>
<th>Other</th>
<th>Sum Alive</th>
<th>CVD</th>
<th>Cancer</th>
<th>Other</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>2621</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2621</td>
<td>1882</td>
<td>0</td>
<td>0</td>
<td>1882</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>276</td>
<td>0</td>
<td>276</td>
<td>0</td>
<td>197</td>
<td>0</td>
<td>0</td>
<td>197</td>
</tr>
<tr>
<td>CVD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>174</td>
<td>0</td>
<td>0</td>
<td>174</td>
<td>0</td>
<td>174</td>
</tr>
<tr>
<td>Cancer</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Acute DM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Lung</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>GI</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>45</td>
</tr>
<tr>
<td>Infection</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>52</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>Accidents</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>172</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>172</td>
</tr>
<tr>
<td>Sum</td>
<td>2621</td>
<td>276</td>
<td>174</td>
<td>352</td>
<td>3423</td>
<td>1882</td>
<td>197</td>
<td>111</td>
<td>2421</td>
</tr>
</tbody>
</table>

> save( C1, file="./data/T2CoD.Rda" )
```
Once we have groomed the L1 dataset we can start by making the baseline table (table 1):

```
> tab1 <- NArray(
+   list( sex = levels(L1$sex),
+         c("N", "Age", "Age-IQR",
+           "DM dur", "DMdur-IQR",
+           "FU time", "FUtime-IQR",
+           "Nephropathy", "Neuropathy",
+           "Retinopathy", "CVD",
+           "Deaths"),
+         c("N", "/%sd") )
> str( tab1 )
logi [1:2, 1:12, 1:2] NA NA NA NA NA NA ...
- attr(*, "dimnames")=List of 3
..$ sex: chr [1:2] "M" "F"
..$ : chr [1:12] "N" "Age" "Age-IQR" "DM dur" ...
..$ : chr [1:2] "N" "/%sd"
> tab1[,"N","N"] <- with( L1, table(sex) )
> tab1[,"M","N","%/sd"] <- tab1[,"M","N","N"]/(tab1[,"M","N","N"]+tab1[,"F","N","N"])*100
> tab1[,"Age","N"] <- with( L1, tapply(entry+0.5-dobth,sex,median) )
> tab1[,"Age","%/sd"] <- with( L1, tapply(entry+0.5-dobth,sex,sd) )
> tab1[,"Age-IQR","1"] <- with( L1, tapply(entry+0.5-dobth,sex,quantile,probs=1/4) )
> tab1[,"Age-IQR","2"] <- with( L1, tapply(entry+0.5-dobth,sex,quantile,probs=3/4) )
> tab1[,"DM dur","N"] <- with( L1, tapply(entry+0.5-dodm,sex,median) )
> tab1[,"DMdur-IQR","1"] <- with( L1, tapply(entry+0.5-dodm,sex,quantile,probs=1/4) )
> tab1[,"DMdur-IQR","2"] <- with( L1, tapply(entry+0.5-dodm,sex,quantile,probs=3/4) )
> tab1[,"FU time","N"] <- with( L1, tapply(lx.dur,sex,median) )
> tab1[,"FU time","%/sd"] <- with( L1, tapply(lx.dur,sex,sd) )
> tab1[,"FUtime-IQR","1"] <- with( L1, tapply(lx.dur,sex,quantile,probs=1/4) )
> tab1[,"FUtime-IQR","2"] <- with( L1, tapply(lx.dur,sex,quantile,probs=3/4) )
> tab1[,"Nephropathy","N"] <- with( L1, table((entry+0.5)>donef,sex )["TRUE",] )
> tab1[,"Neuropathy","N"] <- with( L1, table((entry+0.5)>doneu,sex )["TRUE",] )
> tab1[,"Retinopathy","N"] <- with( L1, table((entry+0.5)>dodr,sex )["TRUE",] )
> tab1[,"CVD","N"] <- with( L1, table((entry+0.5)>docvd,sex )["TRUE",] )
> tab1[,"Deaths","N"] <- with( L1, table(lex.Xst=="Dead",sex )["TRUE",] )
> tab1[,8:12,"%/sd"] <- tab1[,8:12,"N"] / tab1[,"N",rep(1,5)] * 100
> round( ftable(tab1,col.vars=c(1,3)), 1 )
```

```
 N  3423.0  58.6  2421.0  41.4
 Age  61.0  12.0  62.5  14.1
 Age-IQR  52.9  68.6  52.8  72.1
 DM dur  7.3  7.4  7.8  7.9
 DMdur-IQR  2.6  13.1  3.1  13.6
 FU time  6.0  3.0  6.1  3.0
 FUtime-IQR  2.8  9.0  3.1  9.0
 Nephropathy  1173.0  34.3  566.0  23.4
 Neuropathy  1613.0  47.1  963.0  39.8
 Retinopathy  1426.0  41.7  983.0  40.6
 CVD  2423.0  70.8  1630.0  67.3
 Deaths  802.0  23.4  599.0  22.3
```

1.5 Base tables

The state of DN is defined as presence of complications 180 days after entry:
Data preparation

```r
> round( ftable(tab1, col.vars=c(1,3)), 0 )
```

<table>
<thead>
<tr>
<th>sex</th>
<th>M</th>
<th>F</th>
<th>N %/sd</th>
<th>N %/sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>3423</td>
<td>2421</td>
<td>59</td>
<td>41</td>
</tr>
<tr>
<td>Age</td>
<td>61</td>
<td>62</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Age-IQR</td>
<td>53</td>
<td>53</td>
<td>69</td>
<td>72</td>
</tr>
<tr>
<td>DM dur</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>DMdur-IQR</td>
<td>3</td>
<td>3</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>FU time</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>FUtime-IQR</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Nephropathy</td>
<td>1173</td>
<td>566</td>
<td>34</td>
<td>23</td>
</tr>
<tr>
<td>Neuropathy</td>
<td>1613</td>
<td>563</td>
<td>47</td>
<td>40</td>
</tr>
<tr>
<td>Retinopathy</td>
<td>1426</td>
<td>983</td>
<td>42</td>
<td>41</td>
</tr>
<tr>
<td>CVD</td>
<td>2423</td>
<td>1630</td>
<td>71</td>
<td>67</td>
</tr>
<tr>
<td>Deaths</td>
<td>802</td>
<td>539</td>
<td>23</td>
<td>22</td>
</tr>
</tbody>
</table>
Chapter 2

Mortality by cause of death

In order to model the mortality rates properly, we split the follow-up in smaller intervals (in this case along the calendar time scale):

```r
> load( file="./data/T2CoD.Rda" )

> S1 <- splitLexis( C1, time.scale="per", breaks=seq(1980,2015,1/4) )
> summary( S1 )
```

Transitions:

<table>
<thead>
<tr>
<th>To</th>
<th>From</th>
<th>Alive</th>
<th>CVD</th>
<th>Cancer</th>
<th>Other</th>
<th>Records</th>
<th>Events</th>
<th>Risk time</th>
<th>Persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alive</td>
<td>132598</td>
<td>473</td>
<td>285</td>
<td>583</td>
<td>133939</td>
<td>1341</td>
<td>32913.46</td>
<td>5844</td>
<td></td>
</tr>
</tbody>
</table>

First we illustrate the number of deaths by each cause and by type of diabetes:

```r
> boxes.Lexis( C1,list( x=c(20,80,80,80), y=c(50,90,50,10) ),
+ scale.R=100, DR.sep=c("\n","% / yr"), pos.arr=0.3 )
```

Moreover, we show how the distribution of age and diabetes duration is over the follow-up in a Lexis diagram. We compute the height and width of the graph in order to get proper Lexis diagrams:

```r
> xl <- c(0,60)
> yl <- c(0,100)
> ypi <- 16
> c( diff(xl)/ypi,
+     diff(yl)/ypi )+1

[1] 4.75 7.25
```

```r
> par( mai=c(3,3,1,1)/4, mgp=c(3,1,0)/1.6, las=1 )
> plot ( S1, time.scale=2:1,
+     col=gray(0.6), grid=seq(5,100,5), lty.grid=1, col.grid=gray(0.9),
+     xlim=c(0,60), ylim=c(0,100), xaxs="i", yaxs="i",
+     xlab="Diabetes duration", ylab="Age" )
> points( S1, pch=16, cex=0.5,
+     col=c("transparent",rainbow(3))[S1$lex.Xst] )
> rect( 48,0,60,17, col="white", border="lightgray" )
> text( rep(58.5, 1:4*3.5, c(levels(S1$lex.Xst)[2:4],"C.o.D."),
+     col=c(rainbow(3),"gray"), adj=1, cex=0.9, font=2 )
> box()
```
Figure 2.1: Person-years, deaths and mortality rates by cause of death in the SDC T2 patient population.
Figure 2.2: Distribution of follow-up and deaths (by cause) for T2. Although not visible directly, no person in this plot has a life-line (total follow-up) of more than 7 years, since the earliest entry is 1.1.2002, and the latest exit is 31.12.2008.
2.1 Statistical analysis

2.1.1 Setup

We first set up the modeling parameters for the age, period and duration effects:

```r
> n.pr <- 100
> (a.kn <- with(subset(S1, lex.Xst!="Alive"),
+ quantile(age+lex.dur,probs=c(1,3,5,7,9)/10) ) )
              10%      30%      50%      70%      90%
59.641342    67.638604  74.168378  79.989049  85.806982
> a.pr <- seq(40,95,,n.pr)
> a.ct <- Ns(a.pr, knots=a.kn)
> pref <- 2010
> (p.kn <- with(subset(S1, lex.Xst!="Alive"),
+ quantile(per+lex.dur,probs=c(1,5,9)/10) ) )
              10%    50%     90%
> p.pr <- seq(2002,2011,,n.pr)
> p.ct <- Ns(p.pr, knots=p.kn)
> p.rf <- Ns(rep(pref,n.pr), knots=p.kn)
> dref <- 10
> (d.kn <- with(subset(S1, lex.Xst!="Alive"),
+ c(0,quantile(Ddur+lex.dur,probs=1:2/3,na.rm=TRUE)) ) )
              33.333333%  66.666667%
0.000000 11.363240 19.399553
> d.pr <- seq(0,40,,n.pr)
> d.ct <- Ns(d.pr, knots=d.kn)
> d.rf <- Ns(rep(dref,n.pr), knots=d.kn)
```

Finally we can model the cause-specific mortality rates as a function of age and calendar time, and plot the rates and the RRs. But we first set up an array to hold the predicted rates and RRs:

```r
> res <- NArray(list(pred=c("Ainc","PRR"),
+ cod=levels(C1$lex.Xst)[-1],
+ sex=levels(S1$sex),
+ x=1:n.pr,
+ what=c("Est","lo","hi") ) )
> lin <- NArray(c(dimnames(res)[2:3],
+ list(what=c("P(lin)","RR/year","lo","hi","P(null)")) ) )
> str( res )
logi [1:2, 1:3, 1:2, 1:100, 1:3] NA NA NA NA NA NA ...
- attr(*, "dimnames")=List of 5
  ..$ pred: chr [1:2] "Ainc" "PRR"
  ..$ cod : chr [1:3] "CVD" "Cancer" "Other"
  ..$ sex : chr [1:2] "M" "F"
  ..$ x : chr [1:100] "1" "2" "3" "4" ...
  ..$ what: chr [1:3] "Est" "lo" "hi"
> str( lin )
logi [1:3, 1:2, 1:5] NA NA NA NA NA NA ...
- attr(*, "dimnames")=List of 3
  ..$ cod : chr [1:3] "CVD" "Cancer" "Other"
  ..$ sex : chr [1:2] "M" "F"
  ..$ what: chr [1:5] "P(lin)" "RR/year" "lo" "hi" ...
```
2.2 Age and date of follow-up

Then we fit models for all combinations of sex and diabetes type:

```r
> system.time(
+ for( sx in dimnames(res)[["sex"]])
+ for( cd in dimnames(res)[["cod"]])
+ {
+  map <- glm( (lex.Xst == cd) ~ Ns( per, knots=p.kn )
+    + Ns( age, knots=a.kn ),
+    offset = log(lex.dur/100),
+    family = poisson,
+    data = subset( S1, sex==sx) )
+  mal <- update( map, . ~ . - Ns( per, knots=p.kn ) + per )
+  lin[cd,sx,] <- c( anova( map, mal, test="Chisq" )[2,"Pr(>Chi)"],
+                  ci.lin( mal, subset="per", Exp=TRUE )[c(5:7,4)] )
+  res["Ainc",cd,sx,,] <- ci.exp( map, ctr.mat=cbind(1,p.rf,a.ct) )
+  res["PRR" ,cd,sx,,] <- ci.exp( map, subset="per", ctr.mat=p.ct-p.rf )
+  }
+
user  system elapsed
31.45 2.81 35.61
```

```r
> round( ftable( lin ), 3 )
```

<table>
<thead>
<tr>
<th>what</th>
<th>P(lin) RR/year</th>
<th>lo</th>
<th>hi</th>
<th>P(null)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cod</td>
<td>sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVD</td>
<td>M</td>
<td>0.541</td>
<td>0.902</td>
<td>0.861</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.356</td>
<td>0.910</td>
<td>0.961</td>
</tr>
<tr>
<td>Cancer</td>
<td>M</td>
<td>0.621</td>
<td>0.981</td>
<td>0.926</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.743</td>
<td>1.006</td>
<td>0.935</td>
</tr>
<tr>
<td>Other</td>
<td>M</td>
<td>0.325</td>
<td>0.959</td>
<td>0.920</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.058</td>
<td>0.990</td>
<td>0.941</td>
</tr>
</tbody>
</table>

```r
> round( ftable( (lin[,2:4]-1)*100 ), 1 )
```

<table>
<thead>
<tr>
<th>what RR/year</th>
<th>lo</th>
<th>hi</th>
</tr>
</thead>
<tbody>
<tr>
<td>cod</td>
<td>sex</td>
<td></td>
</tr>
<tr>
<td>CVD</td>
<td>M</td>
<td>-9.8</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>-9.0</td>
</tr>
<tr>
<td>Cancer</td>
<td>M</td>
<td>-1.9</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.6</td>
</tr>
<tr>
<td>Other</td>
<td>M</td>
<td>-4.1</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>-1.0</td>
</tr>
</tbody>
</table>
```

We see that there are no signs of non-linear decrease in mortality rates for any of the causes. Also we see that the decrease in mortality is significant for CVD, some 9–10%/year, not detectable for cancer, and for other causes only a borderline significant decrease of some 4%/year, but only for men. Thus it seems that the major change in mortality rates among the T2D patients is for CVD mortality.

We can now plot the age-effects:
Mortality by cause of death

It is pretty clear from figure 2.3 that the decrease in mortality is mainly for CVD mortality; actually data are compatible with models for each of the cause-specific mortalities with a constant annual change in mortality from each of the causes, but only for CVD this annual change is significantly different from 0.
Figure 2.3: Age-specific mortality and the change of this over calendar time for T2 diabetes patients. Red: F; blue: M, thin lines: 95% c.i.
2.3 Including diabetes duration

It would be prudent to control for diabetes duration too. This is simply done by replicating
the code above, starting with the array to collect effects, and the array to collect the
p-values for linear effects and the estimates of those:

```r
> dnam <- dimnames(res)
> dnam["pred"] <- c(dnam["pred"], "DRR")
> resx <- NArray(dnam)
> str(resx)
```

```
logi [1:3, 1:3, 1:2, 1:100, 1:3] NA NA NA NA NA NA ... 
- attr(*, "dimnames")=List of 5
  ..$ pred: chr [1:3] "Ainc" "PRR" "DRR"
  ..$ cod : chr [1:3] "CVD" "Cancer" "Other"
  ..$ sex : chr [1:2] "M" "F"
  ..$ x : chr [1:100] "1" "2" "3" "4" ...
  ..$ what: chr [1:3] "Est" "lo" "hi"
```

```r
> linx <- NArray(c(dimnames(resx)[2:3],
+ list(eff=c("PRR", "DRR", "DoDM"),
+ what=c("P(lin)", "RR/year", "lo", "hi") ))
> str(linx)
```

```
logi [1:3, 1:2, 1:3, 1:4] NA NA NA NA NA NA ... 
- attr(*, "dimnames")=List of 4
  ..$ cod : chr [1:3] "CVD" "Cancer" "Other"
  ..$ sex : chr [1:2] "M" "F"
  ..$ eff : chr [1:3] "PRR" "DRR" "DoDM"
  ..$ what: chr [1:4] "P(lin)" "RR/year" "lo" "hi"
```

We also want to extract the linear effect of age at diagnosis, so we set up a separate array
to hold these values. Then we fit models for all combinations of sex and cause of death:

```r
> system.time(
+ for( cd in dimnames(resx)["cod"] )
+ for( sx in dimnames(resx)["sex"] )
+ {
+  + mapD <- glm( (lex.Xst == cd) ~ Ns( per, knots=p.kn ) + I(age-Ddur)
+  +   + Ns( age, knots=a.kn )
+  +   + Ns( Ddur, knots=d.kn ),
+  +   offset = log(lex.dur/100),
+  +   family = poisson,
+  +   data = subset( S1, sex==sx ) )
+  + mapd <- update( mapD, . ~ . - I(age-Ddur) )
+  + mapl <- update( mapd, . ~ . - Ns( per, knots=p.kn ) + per )
+  + madl <- update( mapd, . ~ . - Ns( Ddur, knots=d.kn ) + Ddur )
+  + linx[cd,sx,,] <- cbind( anova( mapl, mapd, madl, test="Chisq" )[c(2:3,1),"Pr(>Chi)"],
+  +   + rbind( ci.exp( mapl, subset="per" ),
+  +   +   + ci.exp( madl, subset="Ddur" ),
+  +   +   + ci.exp( mapD, subset="I\("")") )
+  + resx["Ainc",cd,sx,,] <- ci.exp( mapd, ctr.mat=cbind(1,p.rf,a.ct,d.rf) )
+  + resx["PRR",cd,sx,,] <- ci.exp( mapd, subset="per", ctr.mat=p.ct-p.rf )
+  + resx["DRR",cd,sx,,] <- ci.exp( mapd, subset="dur", ctr.mat=d.ct-d.rf )
+  + }
```

```
user  system  elapsed
63.88   1.42   67.33
```
We can now inspect the tests for linearity of period and duration effects as well as the estimates of the slope of the linear effects under the null:

```r
> round( ftable( linx, col.vars=c(2,4) ), 3 )
```

<table>
<thead>
<tr>
<th>sex</th>
<th>cod</th>
<th>eff</th>
<th>P(lin) RR/year</th>
<th>lo</th>
<th>hi</th>
<th>P(lin) RR/year</th>
<th>lo</th>
<th>hi</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>CVD</td>
<td>PRR</td>
<td>0.466</td>
<td>0.897</td>
<td>0.857</td>
<td>0.940</td>
<td>0.856</td>
<td>0.986</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.904</td>
<td>0.857</td>
<td></td>
<td>0.956</td>
<td>0.955</td>
</tr>
<tr>
<td></td>
<td>DRR</td>
<td></td>
<td>0.498</td>
<td>1.033</td>
<td>1.019</td>
<td>1.047</td>
<td>0.989</td>
<td>1.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.012</td>
<td>1.014</td>
<td></td>
<td>1.042</td>
<td>1.042</td>
</tr>
<tr>
<td></td>
<td>DoDM</td>
<td>NA</td>
<td>0.964</td>
<td>0.945</td>
<td>0.983</td>
<td>NA</td>
<td>0.974</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>Cancer</td>
<td>PRR</td>
<td>0.364</td>
<td>0.983</td>
<td>0.927</td>
<td>1.042</td>
<td>0.695</td>
<td>1.014</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.014</td>
<td>0.942</td>
<td></td>
<td>1.092</td>
<td>1.092</td>
</tr>
<tr>
<td></td>
<td>DRR</td>
<td></td>
<td>0.342</td>
<td>0.996</td>
<td>0.977</td>
<td>1.015</td>
<td>0.272</td>
<td>0.981</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.958</td>
<td>1.004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DoDM</td>
<td>NA</td>
<td>1.009</td>
<td>0.989</td>
<td>1.030</td>
<td>NA</td>
<td>1.026</td>
<td>1.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>Other</td>
<td>PRR</td>
<td>0.411</td>
<td>0.952</td>
<td>0.914</td>
<td>1.049</td>
<td>0.247</td>
<td>1.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.011</td>
<td>0.997</td>
<td></td>
<td>1.025</td>
<td>1.025</td>
</tr>
<tr>
<td></td>
<td>DRR</td>
<td></td>
<td>0.052</td>
<td>1.037</td>
<td>1.025</td>
<td>1.049</td>
<td>0.247</td>
<td>1.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.011</td>
<td>0.997</td>
<td></td>
<td>1.025</td>
<td>1.025</td>
</tr>
<tr>
<td></td>
<td>DoDM</td>
<td>NA</td>
<td>0.952</td>
<td>0.935</td>
<td>0.970</td>
<td>NA</td>
<td>0.980</td>
<td>0.960</td>
</tr>
</tbody>
</table>

We can translate these RRs to annual changes in mortality from different causes in units of percentage per year:

```r
> round( ftable( (linx[,,,2:4]-1)*100, col.vars=3:4 ), 1 )
```

<table>
<thead>
<tr>
<th>eff</th>
<th>cod</th>
<th>sex</th>
<th>RR/year</th>
<th>lo</th>
<th>hi</th>
<th>RR/year</th>
<th>lo</th>
<th>hi</th>
<th>RR/year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CVD</td>
<td>M</td>
<td>-10.3</td>
<td>-14.3</td>
<td>-6.0</td>
<td>3.3</td>
<td>1.9</td>
<td>4.7</td>
<td>-3.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>-9.6</td>
<td>-14.4</td>
<td>-4.5</td>
<td>2.7</td>
<td>1.2</td>
<td>4.2</td>
<td>-2.6</td>
</tr>
<tr>
<td></td>
<td>Cancer</td>
<td>M</td>
<td>-1.7</td>
<td>-7.3</td>
<td>4.2</td>
<td>-0.4</td>
<td>-2.3</td>
<td>1.5</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>1.4</td>
<td>-5.8</td>
<td>9.2</td>
<td>1.9</td>
<td>-4.2</td>
<td>0.4</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>M</td>
<td>-4.8</td>
<td>-8.6</td>
<td>-0.8</td>
<td>3.7</td>
<td>2.5</td>
<td>4.9</td>
<td>-4.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>-1.5</td>
<td>-6.4</td>
<td>-3.6</td>
<td>1.1</td>
<td>-0.3</td>
<td>2.5</td>
<td>-2.0</td>
</tr>
</tbody>
</table>

Again it appears that there is not much evidence against (log-)linear effects of calendar time and duration, and that the major effect is for CVD +3%/year by duration, −10%/year by calendar time. There is no effect for Cancer, and for Other causes there seems to be a small effect by duration, in the vicinity of 4%/year, but only for men.

```r
cod <- c(1,1000)/30
dimnames(resx)[["cod"]] <- cod
par( mcol=c(3,3), mar=c(3,0,1,0), mgp=c(3,1,0)/1.6, oma=c(0,5,2,1), las=1 )
for( cd in dimnames(resx)[["cod"]] ) {
  matplot( a.pr, a.pr, type="n", log="y", ylim=mlim, yaxt="n",
           xlab="Age", las=1 )
  abline( v=seq(0,100,5), h=outer(1:9,10^(-2:2)), col=gray(0.8) )
  for( sx in 1:2 )
    matlines( a.pr, resx["Ainc",cd,sx,,],
      lwd=3, col=c("blue","red")[sx] )
  mtext( cd, line=0.5, side=3, outer=FALSE )
  if( cd==dimnames(resx)[["cod"]] )
    { axis( side=2 )
      mtext( "Mortality at 1 Jan 2010 (%/year)",
              line=3, side=2, outer=FALSE, las=0 )
    }
  matplot( p.pr, p.pr, type="n", log="y", ylim=rlim, yaxt="n",
            xlab="Date of follow-up", las=1 )
  abline( y=2000+1:15, h=outer(1:9,10^(-2:2)), col=gray(0.8) )
  for( sx in 1:2 )
```
Mortality by cause of death

2.3 Including diabetes duration

```r
+ matlines( p.pr, resx["PRR",cd,sx,,],
  + lwd=c(3,1,1), lty=1, col=c("blue","red")[sx] )
+ abline( h=1 )
+ points( pref,1, pch=16, col="limegreen" )
+ points( dref ,1, pch=1, lws=2 )
+ if( cd==dimnames(resx)[["cod"]][1] )
  + {
    + axis( side=2 )
    + mtext( "Mortality RR", line=3, side=2, outer=FALSE, las=0 )
  + }
+ matplot( d.pr, d.pr, type="n", log="y", ylim=rlim, las=1, yaxt="n",
  + xlab="Duration of diabetes", ylab="" )
+ abline( v=seq(0,100,5), h=outer(1:9,10^(-2:2)), col=gray(0.8) )
+ for( sx in 1:2 )
  + matlines( d.pr, resx["DRR",cd,sx,,],
    + lwd=c(3,1,1), lty=1, col=c("blue","red")[sx] )
  + abline( h=1 )
+ points( dref ,1, pch=16, col="limegreen" )
+ points( dref ,1, pch=1, lws=2 )
+ if( cd==dimnames(resx)[["cod"]][1] )
  + {
    + axis( side=2 )
    + mtext( "Mortality RR", line=3, side=2, outer=FALSE, las=0 )
  + }
+ }
+ }
```
Figure 2.4: Age-specific mortality and the change of this over calendar time for T2 patients. Red: F; blue: M; thin lines: 95% c.i.
Chapter 3

All-cause mortality by complication status

We start by making a tabular overview of how patients go through the study, that is how many have nephropathy at entry, and exit, subdivided by exit status and sex:

```r
> load( file="./data/T2L1.Rda" )
> load( file="./data/T2CoD.Rda" )
> with( C1, ftable( addmargins( table( sex,
+    eDN=!is.na(donef) & entry | is.na(donef)),
+    lex.Xst,
+    xD=!is.na(donef) ),
+    margin = c(1,4) ),
+    col.vars=4:3 )
```

<table>
<thead>
<tr>
<th>xD</th>
<th>FALSE</th>
<th>TRUE</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>lex.Xst</td>
<td>Alive</td>
<td>CVD</td>
<td>Cancer</td>
</tr>
<tr>
<td>sex</td>
<td>eDN</td>
<td>Alive</td>
<td>CVD</td>
</tr>
<tr>
<td>M</td>
<td>FALSE</td>
<td>1658</td>
<td>112</td>
</tr>
<tr>
<td>TRUE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>FALSE</td>
<td>1415</td>
<td>94</td>
</tr>
<tr>
<td>TRUE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sum</td>
<td>FALSE</td>
<td>3073</td>
<td>206</td>
</tr>
<tr>
<td>TRUE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

3.1 Setup

We can now introduce the time-dependent nephropathy status ("DN"):

```r
> C2 <- cutLexis( C1, cut = C1$donef, 
+    timescale = "per",
+    new.state = "DN",
+    new.scale = "DNdur",
+    precursor.states = "Alive" )
> summary( C2, by="sex", scale=100 )
```

$M

Transitions:
From Alive DN CVD Cancer Other Records: Events: Risk time: Persons:
Alive 1691 285 112 81 131 2300 609 115.59 2300

29
All-cause mortality by complication status

<table>
<thead>
<tr>
<th></th>
<th>DN</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>930</td>
<td>1691</td>
</tr>
<tr>
<td>93</td>
<td>221</td>
<td>1408</td>
</tr>
<tr>
<td>1408</td>
<td>75.28</td>
<td>1408</td>
</tr>
<tr>
<td>1087</td>
<td>190.87</td>
<td>3423</td>
</tr>
</tbody>
</table>

$F$

Transitions:

<table>
<thead>
<tr>
<th>From</th>
<th>Alive</th>
<th>DN</th>
<th>CVD</th>
<th>Cancer</th>
<th>Other</th>
<th>Records:</th>
<th>Events:</th>
<th>Risk time:</th>
<th>Persons:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alive</td>
<td>1431</td>
<td>616</td>
<td>197</td>
<td>111</td>
<td>231</td>
<td>2586</td>
<td>704</td>
<td>138.27</td>
<td>2421</td>
</tr>
<tr>
<td>DN</td>
<td>0</td>
<td>451</td>
<td>103</td>
<td>33</td>
<td>114</td>
<td>701</td>
<td>250</td>
<td>36.87</td>
<td>701</td>
</tr>
<tr>
<td>Sum</td>
<td>1431</td>
<td>616</td>
<td>197</td>
<td>111</td>
<td>231</td>
<td>2586</td>
<td>704</td>
<td>138.27</td>
<td>2421</td>
</tr>
</tbody>
</table>

We now have a Lexis object with follow-up along 4 time-scales, age and calendar time as well as duration of diabetes and duration of DN.

We do a few small calculations to enable plotting of the Lexis diagrams properly:

```r
> ypi <- 14
> ll <- 50
> al <- c(0,100)
> dl <- c(0,60)
We can then set up the plot correctly in a pdf-file:
```
All-cause mortality by complication status

Figure 3.1: Follow-up for T2 patients at Steno. Follow-up after onset of DN is shown in dark gray color, deaths shown as dots.
From the boxes in figure 3.2 we see that cancer mortality rates are not affected by the occurrence of DN, whereas rates of death from CVD and other causes are, but CVD death somewhat more.

Figure 3.2: Rates of DN and mortality rates. Number in the boxes are person-years, numbers on arrow are no. transitions (rates per 100 PY).
3.2 Duration of complications

Now recall that the time scale \( \text{DNdur} \) (duration of complications) is not meaningful for persons with a recorded complication date earlier than 1994.02, except as indication of presence of complications by this date. So the complications duration variable must be set to \texttt{NA} for persons for whom the date of complication onset is unknown or non-existent. But we must also construct an indicator for having complications as of 1994 (although formally this could be constructed on the fly as \texttt{is.na(DNdur) & lex.Cst="DN"})

\begin{verbatim}
> tt <- with( C1, table(round(donef,4)) )
> table( tt )

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>7</th>
<th>8</th>
<th>52</th>
<th>265</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1167</td>
<td>254</td>
<td>34</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

> mdat <- as.numeric( names( tt[tt>50] ) )
> class( mdat ) <- "cal.yr"
> as.Date( mdat )

[1] "1993-10-04" "2005-09-30"

> C2 <- transform( C2, DNdur = ifelse( donef < 1994.02, NA, DNdur ),
+     comp94 = factor( pmax(donef<1994.02,0,na.rm=TRUE),
+     labels=c("None","Prev") ) )
> with( C2, table( lex.Cst, comp94, exclude=NULL ) )

<table>
<thead>
<tr>
<th>comp94</th>
<th>lex.Cst</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>4185</td>
</tr>
<tr>
<td>Prev</td>
<td>0</td>
</tr>
<tr>
<td>DN</td>
<td>1811</td>
</tr>
<tr>
<td>CVD</td>
<td>0</td>
</tr>
<tr>
<td>Cancer</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
</tr>
<tr>
<td>&lt;NA&gt;</td>
<td>0</td>
</tr>
</tbody>
</table>

> tt <- with( C2, addmargins( table(donef,comp94,exclude=NULL) ) )
> print.table( tt[c(1:5,nrow(tt)-4:0),],zero.print=".")

<table>
<thead>
<tr>
<th>donef</th>
<th>None</th>
<th>Prev</th>
<th>&lt;NA&gt;</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992.235455516769</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1993.41546885695</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1993.66461327858</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1993.75496235455</td>
<td>265</td>
<td>.</td>
<td>265</td>
<td>265</td>
</tr>
<tr>
<td>1993.75770020534</td>
<td>.</td>
<td>1</td>
<td>.</td>
<td>1</td>
</tr>
<tr>
<td>2012.34496919918</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2012.37782340862</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2012.41067761807</td>
<td>.</td>
<td>1</td>
<td>.</td>
<td>1</td>
</tr>
<tr>
<td>&lt;NA&gt;</td>
<td>3686</td>
<td>.</td>
<td>3686</td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>5996</td>
<td>298</td>
<td>.</td>
<td>6294</td>
</tr>
</tbody>
</table>
\end{verbatim}
3.3 Splitting follow-up for Poisson analysis

In order to model the mortality rates properly, we split the follow-up in smaller intervals (in this case along the calendar time scale):

```r
> S2 <- splitLexis( C2, time.scale="per", breaks=seq(1980,2015,1/4) )
> summary( C2 )

Transitions:
To
From Alive DN CVD Cancer Other Records: Events: Risk time: Persons:
Alive 3122 450 206 159 248 4185 1063 21698.70 4185
DN 0 1381 267 126 335 2109 728 11214.76 2109
Sum 3122 1831 473 285 583 6294 1791 32913.46 5844
```

```r
> summary( S2 )

Transitions:
To
From Alive DN CVD Cancer Other Records: Events: Risk time: Persons:
Alive 87550 450 206 159 248 88613 1063 21698.70 4185
DN 0 45048 267 126 335 45776 728 11214.76 2109
Sum 87550 45498 473 285 583 134389 1791 32913.46 5844
```

So we observe that the time-splitting has expanded the number of records substantially; from 6294 to 134389. To illustrate how each person contributes to the number of records, we show the records from 3 persons:

```r
> options( digits=5 )
> subset( S2, lex.id %in% c(47,108,125,133), select=c(1:8,10) )
```

<table>
<thead>
<tr>
<th>lex.id</th>
<th>age</th>
<th>Ddur</th>
<th>per</th>
<th>DNdur</th>
<th>lex.dur</th>
<th>lex.Cst</th>
<th>lex.Xst</th>
<th>sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>1300</td>
<td>47</td>
<td>80.060</td>
<td>1.37301</td>
<td>2005.1</td>
<td>NA</td>
<td>0.1028405</td>
<td>Alive</td>
<td>M</td>
</tr>
<tr>
<td>1301</td>
<td>47</td>
<td>80.163</td>
<td>1.47585</td>
<td>2005.2</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>M</td>
</tr>
<tr>
<td>1302</td>
<td>47</td>
<td>80.413</td>
<td>1.72585</td>
<td>2005.5</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>M</td>
</tr>
<tr>
<td>1303</td>
<td>47</td>
<td>80.663</td>
<td>1.97585</td>
<td>2005.8</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>M</td>
</tr>
<tr>
<td>1304</td>
<td>47</td>
<td>80.913</td>
<td>2.22585</td>
<td>2006.0</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>M</td>
</tr>
<tr>
<td>1305</td>
<td>47</td>
<td>81.163</td>
<td>2.47585</td>
<td>2006.2</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>M</td>
</tr>
<tr>
<td>1306</td>
<td>47</td>
<td>81.413</td>
<td>2.72585</td>
<td>2006.5</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>M</td>
</tr>
<tr>
<td>1307</td>
<td>47</td>
<td>81.663</td>
<td>2.97585</td>
<td>2006.8</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>M</td>
</tr>
<tr>
<td>1308</td>
<td>47</td>
<td>81.913</td>
<td>3.22585</td>
<td>2007.0</td>
<td>NA</td>
<td>0.2361396</td>
<td>Alive</td>
<td>M</td>
</tr>
<tr>
<td>2591</td>
<td>108</td>
<td>84.752</td>
<td>14.39061</td>
<td>2002.0</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2592</td>
<td>108</td>
<td>85.002</td>
<td>14.64061</td>
<td>2002.2</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2593</td>
<td>108</td>
<td>85.252</td>
<td>14.89061</td>
<td>2002.5</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2594</td>
<td>108</td>
<td>85.502</td>
<td>15.14061</td>
<td>2002.8</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2595</td>
<td>108</td>
<td>85.752</td>
<td>15.39061</td>
<td>2003.0</td>
<td>NA</td>
<td>0.1813826</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2596</td>
<td>108</td>
<td>85.933</td>
<td>15.57199</td>
<td>2003.2</td>
<td>0.0000000</td>
<td>0.0686174</td>
<td>DN</td>
<td>F</td>
</tr>
<tr>
<td>2597</td>
<td>108</td>
<td>86.002</td>
<td>15.64061</td>
<td>2003.2</td>
<td>0.068617</td>
<td>0.2500000</td>
<td>DN</td>
<td>F</td>
</tr>
<tr>
<td>2598</td>
<td>108</td>
<td>86.252</td>
<td>15.89061</td>
<td>2003.5</td>
<td>0.318617</td>
<td>0.2500000</td>
<td>DN</td>
<td>F</td>
</tr>
<tr>
<td>2599</td>
<td>108</td>
<td>86.502</td>
<td>16.14061</td>
<td>2003.8</td>
<td>0.568617</td>
<td>0.0912047</td>
<td>DN</td>
<td>F</td>
</tr>
<tr>
<td>2914</td>
<td>125</td>
<td>54.672</td>
<td>0.73141</td>
<td>2008.9</td>
<td>NA</td>
<td>0.0800821</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2915</td>
<td>125</td>
<td>54.752</td>
<td>0.81149</td>
<td>2009.0</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2916</td>
<td>125</td>
<td>55.002</td>
<td>1.06149</td>
<td>2009.2</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2917</td>
<td>125</td>
<td>55.252</td>
<td>1.31149</td>
<td>2009.5</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2918</td>
<td>125</td>
<td>55.502</td>
<td>1.56149</td>
<td>2009.8</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2919</td>
<td>125</td>
<td>55.752</td>
<td>1.81149</td>
<td>2010.0</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2920</td>
<td>125</td>
<td>56.002</td>
<td>2.06149</td>
<td>2010.2</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2921</td>
<td>125</td>
<td>56.252</td>
<td>2.31149</td>
<td>2010.5</td>
<td>NA</td>
<td>0.2500000</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>2922</td>
<td>125</td>
<td>56.502</td>
<td>2.56149</td>
<td>2010.8</td>
<td>NA</td>
<td>0.2493155</td>
<td>Alive</td>
<td>F</td>
</tr>
<tr>
<td>3121</td>
<td>133</td>
<td>67.670</td>
<td>29.41918</td>
<td>2002.0</td>
<td>1.917564</td>
<td>0.2500000</td>
<td>DN</td>
<td>F</td>
</tr>
</tbody>
</table>
### 3.4 Analysis of duration variables

Since the duration of DN is unknown for a substantial part of the patients with DN, we must either exclude these patients from analysis, which is not feasible, or include them in the analysis as a special group, that is with an indicator of “unknown complications duration”.

But in order to include these in the model we must decide on a value for the complications duration to assign to observations from this group.

Since the parametrization is constructed so that the c.pr effect is 0 at 0, we should code DNdur to 0 for those with unknown duration of complications, as well as for those with no complications at all.

If we include only an indicator of presence of complications (lex.Cst=="DN") we simply assume that presence of complications increase the mortality by a fixed amount. If we also include the indicator Cprev for those with complications as of 1994, the parameter associated with this will be the extra mortality associated with “early” complications.

Finally, if we add the effect of complications duration DNdur, the coefficients of the complications indicator will be the RR for those with “late” complications at the date of complication onset, and the coefficient of Cprev will be the RR of those with “early” complications (pre-1994) relative to those with “late” complications at the date of complication onset.

```r
S2$DNdur <- ifelse( is.na(S2$DNdur), 0, S2$DNdur )
> summary.data.frame( S2 )
```

<table>
<thead>
<tr>
<th>lex.id</th>
<th>age</th>
<th>Ddur</th>
<th>per</th>
<th>DNdur</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>1</td>
<td>14.29</td>
<td>0.000</td>
<td>2002</td>
</tr>
<tr>
<td>1st Qu.</td>
<td>1445</td>
<td>55.27</td>
<td>6.068</td>
<td>2005</td>
</tr>
<tr>
<td>Median</td>
<td>2912</td>
<td>63.52</td>
<td>10.890</td>
<td>2007</td>
</tr>
<tr>
<td>Mean</td>
<td>2916</td>
<td>62.87</td>
<td>12.233</td>
<td>2007</td>
</tr>
<tr>
<td>3rd Qu.</td>
<td>4379</td>
<td>71.85</td>
<td>16.925</td>
<td>2009</td>
</tr>
<tr>
<td>Max.</td>
<td>5844</td>
<td>100.93</td>
<td>57.397</td>
<td>2011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>lex.dur</th>
<th>lex.Cst</th>
<th>lex.Xst</th>
<th>dodth</th>
<th>sex</th>
<th>dobth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>0.0001711</td>
<td>Alive</td>
<td>88613</td>
<td>Alive</td>
<td>87550</td>
</tr>
<tr>
<td>1st Qu.</td>
<td>0.2500000</td>
<td>DN</td>
<td>45776</td>
<td>DN</td>
<td>45498</td>
</tr>
<tr>
<td>Median</td>
<td>0.2500000</td>
<td>CVD</td>
<td>0</td>
<td>CVD</td>
<td>473</td>
</tr>
<tr>
<td>Mean</td>
<td>0.2449118</td>
<td>Cancer</td>
<td>0</td>
<td>Cancer</td>
<td>285</td>
</tr>
<tr>
<td>3rd Qu.</td>
<td>0.2500000</td>
<td>Other</td>
<td>0</td>
<td>Other</td>
<td>583</td>
</tr>
<tr>
<td>Max.</td>
<td>0.2500000</td>
<td>NA's</td>
<td>112191</td>
<td>Max.</td>
<td>1906</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>dodm</th>
<th>exit</th>
<th>entry</th>
<th>codth</th>
<th>docvd</th>
<th>cvd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>1951</td>
<td>Min.</td>
<td>2002</td>
<td>Min.</td>
<td>2002</td>
</tr>
<tr>
<td>Max.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**All-cause mortality by complication status**

3.4 Analysis of duration variables
Note that it is important that the changing of NAs to 0s for these time-scales is done after
time-splitting to avoid that the 0s are taken as actual time-points and subsequent intervals
counted from that on the Ddur and the DNdur scale.
This is a general phenomenon for timescales that are not known for the entire follow-up
of the patients:

- First we must set the date of complications to an arbitrary (early) data for those that
  are prevalent at entry without known date in order to make sure that the state (in
  casu "Complications") is coded correctly.

- Next the time-scale "duration of complications" must be set to NA for those where the
  duration is unknown (whether because they have no complications or because of the
  complications onset is unknown).

- Then we can split the follow-up time, and keep the coding of the duration time-scale
  NA for those with unknown duration.

- And finally, for analysis purposes we must code the duration variable to a
  non-missing value (in casu 0) in order to include it in the modeling.

- This analysis must also always include the indicator variable of having complications,
  and possibly also an indicator of entering the study with prevalent complications.

In the reporting we can reconstruct the ratio of the group with unknown duration
relative to persons with a given duration of complication. However, when reporting this
and other effects we shall further discuss the problems associated with reporting effects of
several simultaneous time scales.

3.4.1 Spline setup
We set up the modeling parameters for the age, period and duration effects. For
convenience in definition of contrast matrices with reference points, we will use the same
number of rows in all matrices.

Note the convention that current age is denoted by a whereas age at diagnosis of DM is
denoted by A, and similarly for period (p, P) and duration of diabetes (d, D) — the latter at
diagnosis of complications:
All-cause mortality by complication status

3.4 Analysis of duration variables

```r
> n.pr <- 100
> ( a.kn <- with( subset(S2,!(lex.Xst %in% c("Alive","DN"))),
+     quantile(age+lex.dur,probs=c(1,3,5,7,9)/10) ) )

 10%  30%  50%  70%  90%
59.64134 67.63860 74.16838 79.98905 85.80698

> a.pr <- seq(40,90,,n.pr)
> a.ct <- Ns( a.pr, knots=a.kn )
> ( A.kn <- with( subset(S2,!(lex.Xst %in% c("Alive","DN"))),
+     quantile(age-Ddur+lex.dur,probs=c(1,3,5,7,9)/10) ) )

 10%  30%  50%  70%  90%
40.81414 50.92603 57.85626 63.65393 72.55930

> A.pr <- seq(5,75,,n.pr)
> A.ct <- Ns( A.pr, knots=A.kn )
> pref <- 2010
> ( p.kn <- with( subset(S2,!(lex.Xst %in% c("Alive","DN"))),
+     quantile(per+lex.dur,probs=c(1,5,9)/10) ) )

 10%  50%  90%
2003.088 2006.719 2010.130

> p.pr <- seq(2002,2010,,n.pr)
> p.ct <- Ns( p.pr, knots=p.kn )
> p.rf <- Ns( rep(pref,n.pr), knots=p.kn )
> Pref <- 2010
> ( P.kn <- with( subset(S2,!(lex.Xst %in% c("Alive","DN"))),
+     quantile(per-Ddur,probs=c(1,5,9)/10) ) )

 10%  50%  90%

> P.pr <- seq(2000,2010,,n.pr)
> P.ct <- Ns( P.pr, knots=P.kn )
> P.rf <- Ns( rep(Pref,n.pr), knots=P.kn )
> dref <- 10
> ( d.kn <- with( subset(S2,!(lex.Xst %in% c("Alive","DN"))),
+     c(0,quantile(Ddur+lex.dur,probs=1:3/4,na.rm=TRUE)) ) )

 25%  50%  75%
0.000000 9.254377 15.488044 21.66997

> d.pr <- seq(0,50,,n.pr)
> d.ct <- Ns( d.pr, knots=d.kn )
> d.rf <- Ns( rep(dref,n.pr), knots=d.kn )
> Dref <- 10
> ( D.kn <- with( subset(S2,!(lex.Xst %in% c("Alive","DN")) & !is.na(DNdur)),
+     c(0,quantile(Ddur-DNdur+lex.dur,probs=1:4/5,na.rm=TRUE)) ) )
```
We plot the columns of the contrast matrix \( c.ct \) as functions of \( c.pr \) (figure 3.3) to ascertain that the time-scales are actually coded so that the effects are 0 at 0. This is a consequence of explicitly setting the first knot (lower boundary knot) to 0.

> par( mfrow=c(1,2), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6 )
> matplot( d.pr, d.ct, type="l", lwd=2, lty=1, ylim=c(-1,1) )
> abline( v=d.kn )
> matplot( c.pr, c.ct, type="l", lwd=2, lty=1, ylim=c(-1,1) )
> abline( v=c.kn )

Figure 3.3: The columns of the contrast matrices for diabetes and complications durations; note that for both duration codings, the value of all columns is 0 at 0.
3.5 Including population mortality rates

We now further merge in the population mortality data in order to be able to compute SMR using the expected number of cases. Note we use the midpoint of each interval for assigning population rates. Since we only split on calendar time (per), we can only compute the midpoint of each interval on this scale, and then use date of birth to compute the age at this midpoint. Once this is done, we take the integer part in order to have variables that match with those in the population data set. Basically what we are doing is that we for each follow-up interval determine which one-year age and period class the midpoint of the interval belongs to. Note that we use timeBand to determine the interval midpoint; this function gives the midpoint of the interval as it would have been if the person were not truncated, censored or exited to another state.

```r
> data(M.dk)
> M.dk <- transform( M.dk, sex = factor(sex,labels=c("M","F")) )
> S2$P <- timeBand( S2, "per", "middle" )
> S2 <- transform(S2, A=floor(P-dobth), P=floor(P) )
> S2 <- merge(S2, M.dk[,c("A","P","sex","rate")])
> S2 <- transform( S2, E=rate*lex.dur/1000 )
> subset( S2, !(E>0) )
```

```r
[1] sex  P  A  lex.id age  Ddur  per  DNdur lex.dur lex.Cst lex.Xst dodth
[13] dobth dodm  exit  entry codth docvd cvd  dodr  dr  donef  nef  doneu
[25] neu  CoD comp94 rate E
<0 rows> (or 0-length row.names)
```

```r
> S2 <- subset( S2, E>0 )
> str(S2)
```

Classes 'Lexis' and 'data.frame': 134384 obs. of 29 variables:
$ sex : Factor w/ 2 levels "M","F": 2 2 2 2 2 2 2 2 2 2 ...  
$ P : num 2002 2002 2002 2002 2002 ...  
$ A : num 17 17 17 17 17 17 17 17 17 17 ...  
$ lex.id : int 5844 5844 5844 5844 5844 ...  
$ age : num 17.3 17.2 17.9 17 17.8 ...  
$ Ddur : num 4.147 1.975 0.037 3.897 0.27 ...  
$ per : num 2002 2002 2002 2002 2002 ...  
$ DNdur : num 0 0 0 0 0 0 0 0 0 0 ...  
$ lex.dur: num 0.25 0.25 0.25 0.25 0.25 ...  
$ lex.Cst: Factor w/ 5 levels "Alive","DN","CVD",...: 1 1 1 1 1 1 1 1 1 1 ...  
$ lex.Xst: Factor w/ 5 levels "Alive","DN","CVD",...: 1 1 1 1 1 1 1 1 1 1 ...  
$ dodth : num NA NA NA NA NA NA NA NA NA NA ...  
$ dobth : num 1985 1985 1985 1985 1985 ...  
$ exit : num 2012 2012 2012 2012 2012 ...  
$ entry : num 2002 2002 2002 2002 2002 ...  
$ codth : Factor w/ 10 levels ":"","CVD","Cancer",...: 1 1 1 1 1 1 1 1 1 1 ...  
$ docvd : num NA NA NA NA NA NA NA NA NA NA ...  
$ cvd : num NA NA NA NA NA NA NA NA NA NA ...  
$ dodr : num NA NA NA NA NA 2007 ...  
$ dr : num NA NA NA NA NA NA 1 1 1 1 ...  
$ donef : num NA NA NA NA NA NA NA NA NA NA ...  
$ nef : num NA NA NA NA NA NA NA NA NA NA ...  
$ doneu : num NA NA NA NA NA NA NA NA NA NA ...  
$ neu : num NA NA NA NA NA NA NA NA NA NA ...  
$ CoD : Factor w/ 4 levels "Alive","CVD",...: 1 1 1 1 1 1 1 1 1 1 ...  
$ comp94: Factor w/ 2 levels "None","Prev": 1 1 1 1 1 1 1 1 1 1 ...  
$ rate : num 0.252 0.252 0.252 0.252 0.252 ...
3.6 Modeling all cause mortality rates

The basic model must include age, period, diabetes duration and complication status. This will form the basis for exploring extensions before we decide on a model to use for reporting mortality rates and SMR. The model considerations will be made for all cause mortality and SMR and the model structure chosen will be applied to all causes of death.

The general model will be as follows (here fitted for the entire dataset, ignoring sex):

```r
> S3 <- Relevel( S2, list( Dead=3:5 ), first=FALSE )

> m0 <- glm( (lex.Xst=="Dead") ~ Ns( age, knots=a.kn )
+ + Ns( per, knots=p.kn )
+ + Ns( Ddur,knots=d.kn )
+ + I( age*Ddur/100 )
+ + I(lex.Cst=="DN")
+ + comp94,
+ + offset = log(lex.dur/100),
+ family = poisson,
+ data = S3 )

> round( ci.exp( m0 ), 3 )
```
All-cause mortality by complication status

3.6 Modeling all cause mortality rates

\[
\begin{align*}
\exp(\text{Est.}) & \quad 2.5\% \quad 97.5\% \\
(\text{Intercept}) & \quad 1.740 \quad 1.285 \quad 2.355 \\
\text{Ns}(\text{age, knots = a.kn})_1 & \quad 2.339 \quad 1.773 \quad 3.087 \\
\text{Ns}(\text{age, knots = a.kn})_2 & \quad 4.507 \quad 3.542 \quad 5.734 \\
\text{Ns}(\text{age, knots = a.kn})_3 & \quad 11.134 \quad 8.203 \quad 15.111 \\
\text{Ns}(\text{age, knots = a.kn})_4 & \quad 6.057 \quad 4.619 \quad 7.943 \\
\text{Ns}(\text{per, knots = p.kn})_1 & \quad 0.758 \quad 0.594 \quad 0.966 \\
\text{Ns}(\text{per, knots = p.kn})_2 & \quad 0.755 \quad 0.648 \quad 0.878 \\
\text{Ns}(\text{Ddur, knots = d.kn})_1 & \quad 1.803 \quad 1.001 \quad 3.246 \\
\text{Ns}(\text{Ddur, knots = d.kn})_2 & \quad 2.400 \quad 0.658 \quad 8.749 \\
\text{Ns}(\text{Ddur, knots = d.kn})_3 & \quad 2.263 \quad 1.050 \quad 4.877 \\
\text{I}(\text{age} \times \text{Ddur}/100) & \quad 0.956 \quad 0.899 \quad 1.016 \\
\text{I}(\text{lex.Cst == "DN"}) & \quad \text{TRUE} \quad 1.862 \quad 1.657 \quad 2.091 \\
\text{comp94Prev} & \quad 1.128 \quad 0.939 \quad 1.354
\end{align*}
\]

Patients with complications already in 1994 are in this model assumed to have a different mortality from those seeing complications later. Moreover we will explore whether there are different period effects in patients with and without complications.

3.6.1 Timescale selection

First we will explore whether using current age / age at diagnosis or current data / date at diagnosis gives the better description of the mortality rates.

For more detailed duration analyses we test whether current age or age at diagnosis and whether current date or date of diagnosis gives the better description of rates when diabetes duration and complications status is included in the model.

To create an overview we set up an array classified by sex and type of test. As basis model we use the model with current age (age at follow-up, AoF), date of follow-up (PoF), and a single indicator of complication status.

We test 5 changes to the model:

+\text{AoD} Adding age at diagnosis to the model.

−\text{AoF} Removing age at follow-up from the model with both.

−\text{AoD} Removing age at follow-up and age at diagnosis from the model.

+\text{PoD} Adding date of diagnosis to the model.

−\text{PoF} Removing date of follow-up from the model with both.

−\text{PoD} Removing date of follow-up and date of diagnosis entirely.

−\text{comp94} Removing the indicator of prevalent complications from the base model.

and hence the following lay-out of the table with all the tests:

\[
\text{ta} <- \text{NArray}(\text{list}(\text{cod} = \text{c("AllD","SMR")},
+ \text{sex} = \text{levels(S2$sex)},
+ \text{test} = \text{c("+AoD","-AoF","-AoD",}
+ "+PoD","-PoF","-PoD","-comp94")},
+ \text{what} = \text{c("ChiSq","df","P-val")})
\]

We can now fit the models for all 4 combinations of diabetes type and sex, and put the tests in the array. Note that the naming convention for the models here is that \text{A} and \text{P} refer to age and date of diagnosis of DM, whereas \text{a}, \text{p} and \text{d} refer to current age, date and diabetes duration, respectively:
The conclusion from these p-values is that either age at diagnosis or current age could be used. Likewise there is no possibility to choose between period of follow-up and period of diagnosis, but at least for men some period effect is needed. Finally that there is no evidence of difference between patients with complications present in 1994 and those getting complications later.

Hence the base model with the three timescales seems to provide an adequate description for both sexes, and we shall proceed with this.

### 3.7 Estimation and extraction of the results

When we fit a model we extract the results at the same time, which in this case will be the age-specific rates for persons diagnosed in ages 25, 30, . . . , 75, and followed till age 90. However, even if we use current date in the model, we will make predictions ignoring this
by fixing the value for the date of follow-up to 1.1.2010, our reference point for the calendar time effect.

We should keep in mind that the predictions we make are from a model with three timescales: age, duration of diabetes and calendar time. So we should put in values for the period of follow-up that corresponds to a given date of diagnosis, otherwise we will miss out on the trend along the calendar time axis.

Thus we will extract three things from the model:

1. Age-specific mortality rates for persons diagnosed in various ages.
2. The annual change in mortality, overall and subdivided by complication status.
3. The RR between patients with and without complications as of 1.1.2002 and 1.1.2010.

So we fit separate models for both sexes; note that we also have an outer loop over “All cause” and “SMR” (dimnames(res)["cod"] [1:2]).

However we first set up a data frame to use for prediction of mortality rates as a function of age for select values of age at diagnosis, in order to show how the joint effect of age at follow-up, age at diagnosis and duration of diabetes influence the mortality rates. Note that we put per equal to 2010, as a constant, thus we show the predicted mortality of patients diagnosed at a given age in 2010, under the assumption that calendar time effect remains flat at the 2010 level. This is because we want to make predictions over much longer time-spans than we actually have data for, so including increasing calendar time in the predictions would mean that we would have to accommodate predictions of calendar time effects substantially outside our data.

So we should think of the predictions as counterfactuals, that is, how rates would have been if rates were constant at the 2010 level.

As a technical aside, note that the prediction frame we construct contains rows of NAs, in order to get predictions that also have single NAs in it so that plotting of a set of separate age-curves can be done in one statement.

```r
> make.frame <-
+ function( aD ) # aD is age at DM diagnosis
+ {
+   + a.pr <- seq(10,90,0.5)
+   + df <- data.frame( age = a.pr,
+   +     per = 2010,
+   +     Ddur = a.pr-aD,
+   +     lex.Cst = factor( rep(0,length(a.pr)),
+   +     levels=0:2,
+   +     labels=levels(S3$lex.Cst) ),
+   +     lex.dur = 1000,
+   +     E = 1,
+   +     aD = aD )
+   + rbind( NA, df[df$age>aD,] )
+ }
+ ( f0 <- make.frame(40)[1:6,] )
```

<table>
<thead>
<tr>
<th>age</th>
<th>per</th>
<th>Ddur</th>
<th>lex.Cst</th>
<th>lex.dur</th>
<th>E</th>
<th>aD</th>
</tr>
</thead>
<tbody>
<tr>
<td>62</td>
<td>40.5</td>
<td>2010</td>
<td>Alive</td>
<td>1000</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>63</td>
<td>41.0</td>
<td>2010</td>
<td>Alive</td>
<td>1000</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>64</td>
<td>41.5</td>
<td>2010</td>
<td>Alive</td>
<td>1000</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>65</td>
<td>42.0</td>
<td>2010</td>
<td>Alive</td>
<td>1000</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>66</td>
<td>42.5</td>
<td>2010</td>
<td>Alive</td>
<td>1000</td>
<td>1</td>
<td>40</td>
</tr>
</tbody>
</table>
> new.frame <- f0(NULL,)
> for( da in seq(25,75,5) )
+ new.frame <- rbind( new.frame,
+ make.frame(da) )
> str( new.frame )
'data.frame': 891 obs. of 7 variables:
$data.frame': 891 obs. of 7 variables:
$ age : num NA 25.5 26 26.5 27 27.5 28 28.5 29 29.5 ...
$ per : num NA 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...
$ Ddur : num NA 0.5 1 1.5 2 2.5 3 3.5 4 4.5 ...
$ lex.Cst: Factor w/ 3 levels "Alive","DN","Dead": NA 1 1 1 1 1 1 1 1 1 ...
$ lex.dur: num NA 1000 1000 1000 1000 1000 1000 1000 1000 1000 ...
$ E : num NA 1 1 1 1 1 1 1 1 1 ...
$ aD : num NA 25 25 25 25 25 25 25 25 25 ...
>
# A prediction frame for persons with DN
> DN.frame <- transform( new.frame,
+ lex.Cst=factor( as.integer(lex.Cst)+1,
+ levels=1:3,
+ labels=levels(lex.Cst) ) )

We the set up arrays to hold the resulting mortality predictions, the resulting period effects and the tests for the effect of complications. We also fit a model ignoring the complication status altogether (termed the "Naive" model):

> AMort <- NArray( list( cod = c("All cause","SMR","CVDonly","CVD","Cancer"),
+ mod = c("Final","Main","Int","Naive"),
+ state = c("Alive","DN"),
+ sex = levels(S3$sex),
+ pred = new.frame[,"age"],
+ what = c("Est","lo","hi") ) )
> AgeDM <- NArray( dimnames(AMort)[c(1,4,6)] )
> ComplTt <- NArray( c( dimnames(AMort)[c(1,4)],
+ list( pred = c("Linear DdurxAge interaction",
+ "Linear CxP interaction",
+ "Linear period effect",
+ "No CxP interaction",
+ "Compl 1994 ne Compl",
+ "No Compl effect"),
+ what = c("Chisq","df","Pval") ) ) )
> ComplRR <- NArray( c( dimnames(AMort)[c(1,4)],
+ list( pred = c("Compl 1994 vs. later",
+ "Compl later vs. None",
+ "Compl vs. None",
+ "Change / year",
+ "Compl vs. None (2002)",
+ "Compl vs. None (2010)",
+ "Change / year (Compl)",
+ what = c("Est","lo","hi") ) ) )
>
> p.pr <- seq(2001,2011,,30)
> p.rf <- 2010
> p.CMs <- Ns( p.pr , knots=p.kn ) -
+ Ns( rep(p.rf,length(p.pr)), knots=p.kn )
> CurveRR <- NArray( c( dimnames(AMort)[1:4],
+ list( pred = p.pr,
+ what = c("Est","lo","hi") ) ) )

With these structures in place we can now fit the relevant models, extract the tests and the effect parameters.

Since we want to give the RR between patients with complications and patients without both at 2010 and at 2002, we need a contrast matrix to apply to the parameters of the
model apdcll. If we name the three parameters from the terms I(lex.Cst=="DN") and I(per-2010):I(lex.Cst=="DN") as (γ, β₀, β₁), then γ is the log-RR at 2010, but we additionally want the log-RR in 2002, which is:

\[
\gamma + \beta_1(2002 - 2010) - \left( 0 + \beta_0(2002 - 2010) \right) = \gamma + 8\beta_0 - 8\beta_1
\]

So we construct the contrast matrix to provide this parameter too:

```r
> # Contrast matrix to get the RR both at 2010 and at 2002
> ( CMi <- rbind(c(1,8,-8), diag(3)) )
```

```
[,1] [ ,2] [ ,3]
[1,] 1  8 -8
[2,] 1  0  0
[3,] 0  1  0
[4,] 0  0  1
```

```r
> system.time(
+ for( rs in dimnames(AMort)["cod"][1:2] )
+ for( sx in dimnames(AMort)["sex"][1] )
+ {
+   # rs <- dimnames(AMort)["cod"][1]
+   # sx <- dimnames(AMort)["sex"][1]
+   apdci <- glm( (lex.Xst=="Dead") ~ Ns( age , knots=a.kn )
+     +   + Ns( per , knots=p.kn ) :I(lex.Cst=="DN")
+     +   + I(age*Ddur)
+     +   + I(lex.Cst=="DN"),
+     +   offset = log(if(rs=="SMR") E else lex.dur/100),
+     +   family=poisson,
+     +   data = subset(S3,sex==sx) )
+   apdcs <- update( apdci , . ~ . - I(age*Ddur) )
+   apdcl <- update( apdcs , . ~ . - Ns( per , knots=p.kn ) :I(lex.Cst=="DN")
+     +   + I(per-2010):as.numeric(lex.Cst=="DN") )
+   apdcll <- update( apdcs , . ~ . - Ns( per , knots=p.kn ) :I(lex.Cst=="DN")
+     +   + I(per-2010):I(lex.Cst=="DN") )
+   apdc <- update( apdcs , . ~ . - Ns( per , knots=p.kn ) :I(lex.Cst=="DN")
+     +   + I(per-2010) )
+   apdcp <- update( apdcs , . ~ . - I(lex.Cst=="DN") )
+   apD <- update( apdc , . ~ . - Ns( age , knots=a.kn ) + I(age-Ddur)
+     +   + Ns( per , knots=p.kn )
+     +   + Ns( Ddur, knots=d.kn ) )
+   # Tests:
+   ComplTt[rs,sx,,] <- as.matrix( abs( anova( apdcs,
+     +   apdci,
+     +   apdcl,
+     +   apdcll,
+     +   apdc,
+     +   apdcp,
+     +   apd,
+     +   test="Chisq" )[-c(1,7),c(4,3,5)] )
+   # Age-specific mortality rates
+   zf <- predict( apdc , newdata=new.frame, type="link", se.fit=TRUE )
+   zm <- predict( apdcs , newdata=new.frame, type="link", se.fit=TRUE )
+   zi <- predict( apdcl , newdata=new.frame, type="link", se.fit=TRUE )
+   zn <- predict( apd , newdata=new.frame, type="link", se.fit=TRUE )
+   AMort[rs,"Final","Alive",sx,] <- exp( cbind(zf$fit,zf$se.fit) %*% ci.mat() )
+   AMort[rs,"Main","Alive",sx,] <- exp( cbind(zm$fit,zm$se.fit) %*% ci.mat() )
+   AMort[rs,"Int","Alive",sx,] <- exp( cbind(zi$fit,zi$se.fit) %*% ci.mat() )
)
\[ + \text{AMort}[rs,\text{"Naive"},\text{"Alive"},sx,,] <- \exp(\text{cbind}(zn$fit,zn$se.fit)) \times \text{ci.mat()} \]

\[ + \text{zf} <- \text{predict(} \text{apdc, newdata=DN.frame, type="link", se.fit=TRUE}) \]

\[ + \text{zm} <- \text{predict(} \text{apdcs, newdata=DN.frame, type="link", se.fit=TRUE}) \]

\[ + \text{zi} <- \text{predict(} \text{apdc, newdata=DN.frame, type="link", se.fit=TRUE}) \]

\[ + \text{AMort}[rs,\text{"Final"},\text{"DN"},sx,,] <- \exp(\text{cbind}(zf$fit,zf$se.fit)) \times \text{ci.mat()} \]

\[ + \text{AMort}[rs,\text{"Main"},\text{"DN"},sx,,] <- \exp(\text{cbind}(zm$fit,zm$se.fit)) \times \text{ci.mat()} \]

\[ + \text{AMort}[rs,\text{"Int"},\text{"DN"},sx,,] <- \exp(\text{cbind}(zi$fit,zi$se.fit)) \times \text{ci.mat()} \]

\[ + \text{# Age at diagnosis effects} \]

\[ + \text{AgeDM}[rs,sx,] <- \text{ci.exp(} \text{apD, subset="I"}) \]

\[ + \text{# RR by calendar time} \]

\[ + \text{CurveRR}[rs,\text{"Final"},\text{"Alive"},sx,,] <- \text{ci.exp(} \text{apdc, subset="per"}, \text{ctr.mat=cbind(p.pr-p.rf)}) \]

\[ + \text{CurveRR}[rs,\text{"Main"},\text{"Alive"},sx,,] <- \text{ci.exp(} \text{apdcs, subset="FALSE:Ns"}, \text{ctr.mat=p.CMs}) \]

\[ + \text{CurveRR}[rs,\text{"Naive"},\text{"Alive"},sx,,] <- \text{ci.exp(} \text{apd, subset="per"}, \text{ctr.mat=cbind(p.pr-p.rf)}) \]

\[ + \text{# Complication effects} \]

\[ + \text{ComplRR}[rs,sx,\text{"Compl 1994 vs. later"},] <- \text{ci.exp(} \text{apdcp, subset="Prev"}) \]

\[ + \text{ComplRR}[rs,sx,\text{"Compl later vs. None"},] <- \text{ci.exp(} \text{apdcp, subset="DN"}) \]

\[ + \text{ComplRR}[rs,sx,\text{"Change / year"},] <- \text{ci.exp(} \text{apdc, subset="per"}) \]

\[ + \text{ComplRR}[rs,sx,\text{c("Compl vs. None (2002)",} \]

\[ + \text{"Compl vs. None (2010)",} \]

\[ + \text{"Change / year (None)",} \]

\[ + \text{"Change / year (Compl)"},] <- \text{ci.exp(} \text{apdcll,} \]

\[ + \text{subset="DN"}, \]

\[ + \text{ctr.mat=CMi}) \]

\[ + \} \]

\[ \text{user system elapsed} \]

\[ 82.91 1.15 88.98 \]

> \text{## Just to show the parameters actually extracted in the code above:} \]

\[ > \text{ci.exp(} \text{apdc}) \]

<table>
<thead>
<tr>
<th>Exp(Est.)</th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2.4143201</td>
<td>1.5196521</td>
</tr>
<tr>
<td>(age, knots = a.kn)1</td>
<td>0.4555361</td>
<td>0.2976718</td>
</tr>
<tr>
<td>(age, knots = a.kn)2</td>
<td>0.7772101</td>
<td>0.5889821</td>
</tr>
<tr>
<td>(age, knots = a.kn)3</td>
<td>0.5411388</td>
<td>0.3979321</td>
</tr>
<tr>
<td>(age, knots = a.kn)4</td>
<td>0.5717310</td>
<td>0.4464008</td>
</tr>
<tr>
<td>(Ddur, knots = d.kn)1</td>
<td>0.9580779</td>
<td>0.6972619</td>
</tr>
<tr>
<td>(Ddur, knots = d.kn)2</td>
<td>0.6689865</td>
<td>0.2576499</td>
</tr>
<tr>
<td>(Ddur, knots = d.kn)3</td>
<td>1.0722399</td>
<td>0.8240628</td>
</tr>
<tr>
<td>\text{I(lex.Cst == &quot;DN&quot;)TRUE}</td>
<td>1.9926924</td>
<td>1.6724267</td>
</tr>
<tr>
<td>\text{I(per - 2010)}</td>
<td>0.9862718</td>
<td>0.9540445</td>
</tr>
</tbody>
</table>

> \text{ci.exp(} \text{apdcs})

<table>
<thead>
<tr>
<th>Exp(Est.)</th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2.4248983</td>
<td>1.5084862</td>
</tr>
<tr>
<td>(age, knots = a.kn)1</td>
<td>0.4597599</td>
<td>0.3004526</td>
</tr>
<tr>
<td>(age, knots = a.kn)2</td>
<td>0.7694033</td>
<td>0.2654226</td>
</tr>
<tr>
<td>(age, knots = a.kn)3</td>
<td>0.5367366</td>
<td>0.3946416</td>
</tr>
<tr>
<td>(age, knots = a.kn)4</td>
<td>0.5697035</td>
<td>0.4478888</td>
</tr>
<tr>
<td>(Ddur, knots = d.kn)1</td>
<td>0.9768536</td>
<td>0.7104202</td>
</tr>
<tr>
<td>(Ddur, knots = d.kn)2</td>
<td>0.6875692</td>
<td>0.2654226</td>
</tr>
</tbody>
</table>
All-cause mortality by complication status

Estimation and extraction of the results

\[
\text{Ns(Ddur, knots = d.kn)3} & \quad 1.0875588 \quad 0.8537523 \quad 1.3853948 \\
\text{I(lex.Cst == "DN")TRUE} & \quad 2.0108237 \quad 1.4320838 \quad 2.8218702 \\
\text{I(lex.Cst == "DN")FALSE:Ns(per, knots = p.kn)1} & \quad 1.1987803 \quad 0.6944684 \quad 2.0693157 \\
\text{I(lex.Cst == "DN")TRUE:Ns(per, knots = p.kn)1} & \quad 1.0430612 \quad 0.5983416 \quad 1.8183203 \\
\text{I(lex.Cst == "DN")FALSE:Ns(per, knots = p.kn)2} & \quad 0.6521453 \quad 0.4636892 \quad 0.9171951 \\
\text{I(lex.Cst == "DN")TRUE:Ns(per, knots = p.kn)2} & \quad 1.0309941 \quad 0.7323797 \quad 1.4513631 \\
\]

\[
> \text{ci.exp( apdcp )}
\]

\[
\begin{array}{lcccccc}
\text{exp(Est.)} & 2.5\% & 97.5\% \\
\text{(Intercept)} & 2.4294405 & 1.5282849 & 3.8619640 \\
\text{Ns(age, knots = a.kn)1} & 0.4554513 & 0.2975923 & 0.6970473 \\
\text{Ns(age, knots = a.kn)2} & 0.7836084 & 0.5929695 & 1.0355374 \\
\text{Ns(age, knots = a.kn)3} & 0.5424703 & 0.3989120 & 0.7376915 \\
\text{Ns(age, knots = a.kn)4} & 0.5732549 & 0.4475306 & 0.7342988 \\
\text{Ns(Ddur, knots = d.kn)1} & 0.9549534 & 0.6946931 & 1.3127178 \\
\text{Ns(Ddur, knots = d.kn)2} & 0.6650127 & 0.5714235 & 0.7214915 \\
\text{Ns(Ddur, knots = d.kn)3} & 0.6521453 & 0.4636892 & 0.9171951 \\
\text{I(lex.Cst == "DN")TRUE} & 1.9592696 & 1.6267640 & 2.3597383 \\
\text{I(per - 2010) } & 0.9876584 & 0.9550217 & 1.0214105 \\
\text{comp94Prev} & 1.0938205 & 0.7960679 & 1.5029413 \\
\end{array}
\]

\[
> \text{ci.exp( apd )}
\]

\[
\begin{array}{lcccccc}
\text{exp(Est.)} & 2.5\% & 97.5\% \\
\text{(Intercept)} & 2.5437993 & 1.6039846 & 4.0342748 \\
\text{Ns(age, knots = a.kn)1} & 0.4883967 & 0.3196207 & 0.7462950 \\
\text{Ns(age, knots = a.kn)2} & 0.7625204 & 0.5772803 & 1.0072010 \\
\text{Ns(age, knots = a.kn)3} & 0.5363264 & 0.3945901 & 0.7289744 \\
\text{Ns(age, knots = a.kn)4} & 0.5630533 & 0.4394684 & 0.7213921 \\
\text{Ns(Ddur, knots = d.kn)1} & 1.1564677 & 0.8454647 & 1.5818727 \\
\text{Ns(Ddur, knots = d.kn)2} & 0.9277629 & 0.3623429 & 2.3754956 \\
\text{Ns(Ddur, knots = d.kn)3} & 1.2869219 & 1.0151843 & 1.6313964 \\
\text{I(per - 2010) } & 0.9801094 & 0.9481106 & 1.0131883 \\
\end{array}
\]

\[
> \text{save( AMort, AgeDM, CurveRR, ComplRR, ComplTt, file="./data/T2AllCau.Rda" )}
\]

The average change in mortality by age at onset are:

\[
> \text{round( (ftable(AgeDM[1:2,,])-1)*100, 2 )}
\]

<table>
<thead>
<tr>
<th>cod</th>
<th>sex</th>
<th>what</th>
<th>Est</th>
<th>lo</th>
<th>hi</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cause</td>
<td>M</td>
<td>-3.22</td>
<td>-4.33</td>
<td>-2.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>-1.54</td>
<td>-2.91</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>SMR</td>
<td>M</td>
<td>-3.22</td>
<td>-4.33</td>
<td>-2.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>-1.53</td>
<td>-2.90</td>
<td>-0.14</td>
<td></td>
</tr>
</tbody>
</table>

In order to assess how different these slopes by age at diagnosis are, we need a small function to compute the ratio of the rates from the confidence intervals:
```r
> r2rr <-
+ function( r1, r2 )
+ {
+ l1 <- log( r1[1] )
+ sl1 <- log( r1[3]/r1[2] )/(2*1.96)
+ l2 <- log( r2[1] )
+ sl2 <- log( r2[3]/r2[2] )/(2*1.96)
+ lr <- l1-l2
+ slr <- sqrt( sl1^2 + sl2^2 )
+ res <- c( exp( c(lr,slr) %*% ci.mat() ),
+ 1-pchisq( (lr/slr)^2, 1 ) )
+ names( res ) <- c("RR", "lo", "hi", "P")
+ res
+ }
> round( ftable(AgeDM[1:2,,]-1)*100, 1 )

<table>
<thead>
<tr>
<th>cod</th>
<th>sex</th>
<th>Est</th>
<th>lo</th>
<th>hi</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cause</td>
<td>M</td>
<td>-3.2</td>
<td>-4.3</td>
<td>-2.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>-1.5</td>
<td>-2.9</td>
<td>-0.2</td>
<td></td>
</tr>
<tr>
<td>SMR</td>
<td>M</td>
<td>-3.2</td>
<td>-4.3</td>
<td>-2.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>-1.5</td>
<td>-2.9</td>
<td>-0.1</td>
<td></td>
</tr>
</tbody>
</table>

> round( r2rr( AgeDM["All cause","M",], AgeDM["All cause","F",] ), 3 )

<table>
<thead>
<tr>
<th>RR</th>
<th>lo</th>
<th>hi</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.983</td>
<td>0.965</td>
<td>1.001</td>
<td>0.062</td>
</tr>
</tbody>
</table>

> round( r2rr( AgeDM["SMR","M",], AgeDM["SMR","F",] ), 3 )

<table>
<thead>
<tr>
<th>RR</th>
<th>lo</th>
<th>hi</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.983</td>
<td>0.965</td>
<td>1.001</td>
<td>0.061</td>
</tr>
</tbody>
</table>

Thus we see that the male mortality decreases by 3.2% per year of age at diagnosis, the female only by 1.5% per year, the difference is 1.7% (0.0–3.5), p=0.061, so a tendency of a steeper decrease by age at diagnosis for men than women.

A quick look at the tests for the various subsets of data, and the corresponding estimates of effects:

> str( ComplTt)

```
num [1:5, 1:2, 1:6, 1:3] 1.83 1.78 NA NA NA ...  
attr(*, "dimnames")=List of 4
  ..$ cod : chr [1:5] "All cause" "SMR" "CVDonly" "CVD" ...  
  ..$ sex : chr [1:2] "M" "F"  
  ..$ pred: chr [1:6] "Linear DdurxAge interaction" "Linear CxP interaction" "Linear period effect" ...  
  ..$ what: chr [1:3] "Chisq" "df" "Pval"
```

> round( ftable( ComplTt[1:2,,"Pval"], row.vars=3, col.vars=c(1,2) ), 3 )
The table shows that the models with linear period effect is pretty much the adequate one to report from. There is no duration by age interaction, no complications by period interactions and no difference in mortality or SMR between those that have complications before or after 1994.

Thus the base model (apdc in the above code) has an effect of current age, current duration of diabetes, a fixed complications effect and a (log-)linear change in mortality over calendar time.

In the following, the first 2 estimates are from a model where there is an additional complications effect for those with complication onset before 1994 (that is, with unknown complications duration), the next two estimates are from the base model, and the last 4 estimates are from the model with a separate complications RR and separate calendar time trend between persons with and without complications (that is the model including the non-significant period by complications interaction):

\[
\begin{array}{cccccc}
\text{cod} & \text{All cause SMR} \\
\text{sex} & \text{M} & \text{F} & \text{M} & \text{F} \\
\text{pred} \\
\text{Linear DdurxAge interaction} & 0.176 & 0.964 & 0.183 & 0.956 \\
\text{Linear CxP interaction} & 0.281 & 0.248 & 0.292 & 0.259 \\
\text{Linear period effect} & 0.962 & 0.120 & 0.931 & 0.086 \\
\text{No CxP interaction} & 0.095 & 0.283 & 0.084 & 0.281 \\
\text{Compl 1994 ne Compl} & 0.174 & 0.616 & 0.180 & 0.582 \\
\text{No Compl effect} & 0.000 & 0.000 & 0.000 & 0.000 \\
\end{array}
\]

The general picture from the interaction is that the complications RR in 2010 is slightly above 2, and that there is an annual decrease in mortality of 5% (3–8) for men and a decrease of 3% (0–6) for women, whereas the corresponding decreases in SMR are M: 3% (0–6) and W: 1% (−2–5). The changes are not significantly different between men and women, though. We here make a quick test for the equality of the trends:

\[
\begin{array}{cccccc}
\text{cod} & \text{All cause SMR} \\
\text{sex} & \text{M} & \text{F} & \text{M} & \text{F} \\
\text{pred} \\
\text{Compl 1994 vs. later} & 1.17 & 0.93 & 1.46 & 1.09 & 0.79 & 1.49 & 1.17 & 0.93 & 1.46 & 1.09 & 0.80 & 1.50 \\
\text{Compl later vs. None} & 1.71 & 1.47 & 1.99 & 1.97 & 1.63 & 2.37 & 1.71 & 1.47 & 1.99 & 1.96 & 1.63 & 2.36 \\
\text{Compl vs. None} & 1.76 & 1.52 & 2.03 & 2.00 & 1.68 & 2.38 & 1.76 & 1.52 & 2.03 & 1.99 & 1.67 & 2.37 \\
\text{Change / year} & 0.95 & 0.92 & 0.97 & 0.97 & 0.94 & 1.00 & 0.97 & 0.94 & 1.00 & 0.99 & 0.95 & 1.02 \\
\text{Compl vs. None (2002)} & 1.42 & 1.07 & 1.89 & 1.69 & 1.19 & 2.40 & 1.41 & 1.06 & 1.88 & 1.69 & 1.19 & 2.39 \\
\text{Compl vs. None (2010)} & 2.06 & 1.62 & 2.62 & 2.26 & 1.70 & 3.01 & 2.07 & 1.63 & 2.64 & 2.26 & 1.70 & 3.00 \\
\text{Change / year (None)} & 0.92 & 0.88 & 0.96 & 0.95 & 0.91 & 1.00 & 0.94 & 0.90 & 0.98 & 0.97 & 0.93 & 1.01 \\
\text{Change / year (Compl)} & 0.96 & 0.93 & 1.00 & 0.99 & 0.94 & 1.04 & 0.99 & 0.95 & 1.02 & 1.01 & 0.96 & 1.06 \\
\end{array}
\]
We then transform this to the relevant numbers for the table in the paper, that is, RR at fixed times and % change for the mortality trends.

```r
> cod <- ComplRR[c("All cause","SMR"),,"Change / year",]

cod what Est lo hi
All cause M 0.9452018 0.9198842 0.9712162
 F 0.9672470 0.9355611 1.0000060
SMR M 0.9695112 0.9435926 0.9961417
 F 0.9862718 0.9540445 1.0195876

> round( ci.indep( ComplRR["All cause",,"Change / year",] ), 3 )

RR lo up Pval
1.023 0.960 1.068 0.293

> round( ci.indep( ComplRR["SMR" ,,"Change / year",] ), 3 )

RR lo up Pval
1.017 0.975 1.062 0.433
```

We then transform this to the relevant numbers for the table in the paper, that is, RR at fixed times and % change for the mortality trends.
From this we see that there is a significant decrease of 5.5% per year for male T2D and a 3.3% borderline significant decrease in mortality among women.

With these estimates we can now plot the age-effects for all cause mortality:

```r
> # The age at diagnosis
> pref <- 2010
> pr.A <- as.numeric( dimnames(AMort)[["pred"]])
> agr <- cumsum(is.na(pr.A))
> wh <- agr %in% c(4,7,10) ## Age at dx 40 55 70
> mlim <- c(1,900)/3
> par( mar=c(3,3.5,1,1), mgp=c(3,1,0)/1.6 )
> plot( NA, type="n", log="y", ylim=mlim, xlim=c(40,90),
+ xlab="Age", ylab="", las=1 )
> abline( v=seq(0,100,5), h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
> for( sx in 2:1 ) # sx <- 1
+ matlines( pr.A[wh],
+ cbind( AMort["All cause","Final","Alive",sx,wh,],
+ AMort["All cause","Final","DN",sx,wh,] ),
+ lwd=c(3,1,1), lty=rep(c(1,3),each=3), col=c("blue","red")[sx] )
> mtext( paste("Mortality at 1 Jan", pref, "(per 1000 PY)"), line=2.5, side=2, outer=FALSE )
> box()
```

We can do the same for the corresponding SMR-model, giving the SMR by age for a fixed value of period (1.1.2008):

```r
> rlim <- c(1/2,10)
> par( mar=c(3,3.5,1,1), mgp=c(3,1,0)/1.6 )
> plot( NA, type="n", log="y", ylim=rlim, xlim=c(40,90),
+ xlab="Age", ylab="", las=1 )
> abline( v=seq(0,100,5), h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
> for( sx in 2:1 ) # sx <- 1
+ matlines( as.numeric( dimnames(AMort)[["pred"]])[wh],
+ cbind( AMort["SMR","Final","Alive",sx,wh,],
+ AMort["SMR","Final","DN",sx,wh,] ),
+ lwd=c(3,1,1), lty=rep(c(1,3),each=3), col=c("blue","red")[sx] )
> mtext( paste("SMR at 1 Jan", pref, "(per 1000 PY)"), line=2.5, side=2, outer=FALSE )
> box()
```

We also show the shape of the RR as a function of (current) calendar time, both for the model with and without nephropathy interaction:

```r
> rlim <- c(1/2,5)
> pr.pr <- as.numeric( dimnames(CurveRR)[["pred"]])
> par( mar=c(3,3.5,1,1), mgp=c(3,1,0)/1.6 )
> plot( NA, type="n", log="y", ylim=rlim, xlim=c(2002,2011),
+ xlab="Date of FU", ylab="", las=1 )
> abline( v=2000:2015, h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
```
Figure 3.4: Age-specific all-cause mortality rates without (full lines) and with (broken lines) DN for T2 patients aged 40, 55 and 75 at diagnosis (assumed to be in 2010). Red: F; blue: M; thin lines: 95% c.i.
Figure 3.5: *Age-specific all-cause relative mortality (SMR). T2 patients aged 15, 30, 45 at diagnosis, T2 patients age 40, 55, 70 at diagnosis. Red: F; blue: M; thin lines: 95% c.i.*
To explore the shape of the (non-significant) interactions we plot all the mortality curves for the two models together:

```r
> for( rs in dimnames(AMort)[[1]][1:2] )
  + for( md in dimnames(AMort)[[2]] )
  + {
    + plot( NA, type="n", log="y", ylim=if(rs=="SMR") rlim else mlim,
    + xlim=c(40,90), xlab="Age", ylab="", las=1 )
    + abline( v=seq(0,100,5), h=outer(1:15,10^(-1:2),"*"), col=gray(0.8) )
    + if( rs=="All cause" )
    + mtext( paste( md, "model" ), side=3, line=1 )
    + if( md=="Final" )
    + mtext( c( paste("Mortality at 1 Jan", pref, "(per 1000 PY)"),
    +         paste("SMR 1 Jan", pref ) )[1+(rs=="SMR")],
    +         line=3, side=2 )
    + if(rs=="SMR") abline(h=1)
    + for( sx in 1:2 )
    + matlines( as.numeric( dimnames(AMort)[["pred"] ] ),
    + AMort[rs,md,"Alive",sx,,],
    + lwd=2, lty=c(1,0,0),
    + col=c("blue","red")[sx] )
    + box()
  + }
```

### 3.7.1 Graphs and tables for the paper

First we extract the numbers we need for the table of estimates

```r
> arttab <- ComplRR[1:2,,,
```

```r
> arttab$cod
[1] "All cause" "SMR"
```

```r
> arttab$sex
[1] "M" "F"
```

```r
> arttab$pred
[1] "Compl 1994 vs. later" "Compl later vs. None" "Compl vs. None" "Change / year"
```

```r
> arttab$what
[1] "Est" "lo" "hi"
```
Figure 3.6: Age-specific all-cause mortality and SMR for T2D patients without complications, diagnosed 1.1.2010 in ages 25, 30, . . . 75. The final model have a linear effect of calendar time, the main and the interaction model a 2-parameter-spline effect of calendar time; in all models the calendar time effect is separate for persons with and without complications. The interaction model has an extra 1-parameter (product) interaction between current age and current duration of diabetes. Red: F; blue: M.
> arttab[, , (4, 7, 8)] <- round( ( arttab[, , (4, 7, 8)] - 1 ) * 100, 1 )
> round( ftable( arttab, row.vars = c(3, 1)), 2)

<table>
<thead>
<tr>
<th>pred</th>
<th>sex</th>
<th>cod</th>
<th>M Est</th>
<th>lo</th>
<th>hi</th>
<th>F Est</th>
<th>lo</th>
<th>hi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compl 1994 vs. later All cause</td>
<td>female</td>
<td>1.17</td>
<td>0.93</td>
<td>1.46</td>
<td>1.09</td>
<td>0.79</td>
<td>1.49</td>
<td></td>
</tr>
<tr>
<td>Compl later vs. None All cause</td>
<td>female</td>
<td>1.71</td>
<td>1.47</td>
<td>1.99</td>
<td>1.97</td>
<td>1.63</td>
<td>2.37</td>
<td></td>
</tr>
<tr>
<td>Compl vs. None All cause</td>
<td>female</td>
<td>1.76</td>
<td>1.52</td>
<td>2.03</td>
<td>2.00</td>
<td>1.68</td>
<td>2.38</td>
<td></td>
</tr>
<tr>
<td>Change / year All cause</td>
<td>female</td>
<td>-5.50</td>
<td>-8.00</td>
<td>-2.90</td>
<td>-3.30</td>
<td>-6.40</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Compl vs. None (2002) All cause</td>
<td>female</td>
<td>1.42</td>
<td>1.07</td>
<td>1.89</td>
<td>1.69</td>
<td>1.19</td>
<td>2.40</td>
<td></td>
</tr>
<tr>
<td>Compl vs. None (2010) All cause</td>
<td>female</td>
<td>2.06</td>
<td>1.62</td>
<td>2.62</td>
<td>2.26</td>
<td>1.70</td>
<td>3.01</td>
<td></td>
</tr>
<tr>
<td>Change / year (None) All cause</td>
<td>female</td>
<td>-8.10</td>
<td>-11.90</td>
<td>-4.10</td>
<td>-4.90</td>
<td>-9.20</td>
<td>-0.50</td>
<td></td>
</tr>
<tr>
<td>Change / year (Compl) All cause</td>
<td>female</td>
<td>-3.70</td>
<td>-7.00</td>
<td>-0.30</td>
<td>-1.40</td>
<td>-6.10</td>
<td>3.50</td>
<td></td>
</tr>
</tbody>
</table>

It is the last 6 lines of output that enters as the table in the paper.

We then draw the figure(s) for the paper:
All-cause mortality by complication status

3.7 Estimation and extraction of the results

```r
+ text( floor(pr.A[fs])+1, my, paste(floor(pr.A[fs]), adj=0, col=scol[sx] )
+ )
+ mtext( c("a","b")[sx], side=3, adj=0.01, line=0.2 )
+ mtext( c("Men","Women")[sx], side=3, adj=0.5, line=1 )
+ box()
+ }
+
+ ######################################################################
+ # SMR
+ rlim <- c(1/2,10)
+ for( sx in 1:2 )
+ {
+ plot( NA, type="n", log="y", ylim=rlim, xlim=c(40,90),
+ xlab="Age", ylab="", yaxt="n", las=1 )
+ abline( v=seq(0,100,5), h=outer(1:9,10^-1:2,"*"), col=gray(0.8) )
+ abline( h=1 )
+ matlines( pr.A[wh], AMort["SMR","Final","Alive",sx,wh,cls],
+ lwd=c(3,1,1), lty=1, col=scol[sx] )
+ matlines( pr.A[wh], AMort["SMR","Final","DN",sx,wh,cls],
+ lwd=c(3,1,1), lty="11", lend=1, col=scol[sx] )
+ if( sx==1 )
+ {
+ axis(side=2)
+ mtext( paste("SMR at 1 Jan", pref ),
+ line=2.5, side=2, outer=FALSE, las=0, cex=0.7 )
+ }
+ if( !is.null(ry) )
+ {
+ segments( floor(pr.A[fs]),
+ pmin(ry,apply(AMort["SMR","Final",c("Alive","DN"),sx,fs,1],2,min)),
+ floor(pr.A[fs]),
+ pmax(ry,apply(AMort["SMR","Final",c("Alive","DN"),sx,fs,1],2,max)),
+ col=scol[sx] )
+ text( floor(pr.A[fs])+1, ry, paste(floor(pr.A[fs])), adj=0, col=scol[sx] )
+ }
+ box()
+ mtext( c("c","d")[sx], side=3, adj=0.01, line=0.2 )
+ }
+
+ if( dr ) {
+ ######################################################################
+ # Mortality RR
+ rlim <- c(0.4,4)
+ for( sx in 1:2 )
+ {
+ plot( NA, type="n", log="y", ylim=rlim, xlim=c(2002,2011),
+ xlab="Date of follow-up", ylab="", las=1, yaxt="n" )
+ abline( v=2000:2015, h=outer(1:9,10^-1:2,"*"), col=gray(0.8) )
+ abline( h=1 )
+ matlines( p.pr, CurveRR["All cause","Final","Alive",sx,,cls],
+ lwd=c(3,1,1), lty=1, col=scol[sx] )
+ matlines( p.pr, CurveRR["SMR","Final","Alive",sx,,cls],
+ lwd=c(3,1,1), lty="11", lend=1, col=scol[sx] )
+ if( sx==1 )
+ {
+ axis(side=2)
+ mtext( "SMR ratio (broken) Mortality ratio (full)",
+ line=2.5, side=2, outer=FALSE, las=0, cex=0.7 )
+ }
+ box()
+ mtext( c("e","f")[sx], side=3, adj=0.01, line=0.2 )
+ }
+ }
> f1(aod=c(4,6,8,10))
```
Figure 3.7: Mortality rates (a,b), SMR (c,d) and changes in these (e,f). Patients diagnosed with DM in ages 40, 50, 60 and 70. For mortality rates and SMR, patients without nephropathy are with full lines, patients with nephropathy are with dotted lines. For the changes (e,f), the full lines are changes in mortality, dotted lines changes in SMR. Thin lines indicate 95% confidence intervals throughout.
3.7.2 Results from the naive model

We also show results for the overall mortality and SMR as estimated from the naïve model where nephropathy status is ignored. Hence we only have one set of curves in each graph, referring to the overall mortality regardless of nephropathy status.
Figure 3.8: Mortality rates (a,b), SMR (c,d) and changes in these (e,f). Patients diagnosed with DM in ages 40, 45, . . . , 70. For mortality rates and SMR, patients without nephropathy are with full lines, patients with nephropathy are with dotted lines. For the changes (e,f), the full lines are changes in mortality, dotted lines changes in SMR.
All-cause mortality by complication status

3.7 Estimation and extraction of the results

```r
+ mtext( c("a","b")[sx], side=3, adj=0.01, line=0.2 )
+ mtext( c("Men","Women")[sx], side=3, adj=0.5, line=1 )
+ box()
+ }
+
+ # SMR
+ rlim <- c(1/2,5)
+ for( sx in 1:2 )
+ {
+ plot( NA, type="n", log="y", ylim=rlim, xlim=c(40,90),
+ xlab="Age", ylab="", yaxt="n", las=1 )
+ abline( v=seq(0,100,5), h=outer(c(1.5,1:9),10^(-1:2),"*"), col=gray(0.8) )
+ abline( h=1 )
+ matlines( pr.A[wh], AMort["SMR","Naive","Alive",sx,wh,cls],
+ lwd=c(3,1,1), lty="11", lend=1, col=scol[sx] )
+ if( sx==1 )
+ {
+ axis(side=2)
+ mtext( paste("SMR at 1 Jan", pref ),
+ line=2.5, side=2, outer=FALSE, las=0, cex=0.7 )
+ }
+ if( !is.null(ry) )
+ {
+ segments( floor(pr.A[fs]),
+ pmin(ry,AMort["SMR","Naive","Alive",sx,fs,1]),
+ floor(pr.A[fs]),
+ pmax(ry,AMort["SMR","Naive","Alive",sx,fs,1]),
+ col=scol[sx] )
+ text( floor(pr.A[fs])+1, ry, paste(floor(pr.A[fs])), adj=0, col=scol[sx] )
+ }
+ box()
+ mtext( c("c","d")[sx], side=3, adj=0.01, line=0.2 )
+ }
+
+ if( dr )
+ {
+ # Mortality RR
+ rlim <- c(0.5,5)
+ for( sx in 1:2 )
+ {
+ plot( NA, type="n", log="y", ylim=rlim, xlim=c(2002,2011),
+ xlab="Date of follow-up", ylab="", yaxt="n", las=1 )
+ abline( v=2000:2015, h=outer(c(1.5,1:9),10^(-1:2),"*"), col=gray(0.8) )
+ abline( h=1 )
+ matlines( p.pr, CurveRR["All cause","Naive","Alive",sx,,cls],
+ lwd=c(3,1,1), lty=1, col=scol[sx] )
+ matlines( p.pr, CurveRR["SMR" ,"Naive","Alive",sx,,cls],
+ lwd=c(3,1,1), lty="11", lend=1, col=scol[cls] )
+ points( pref, 1, pch=16, col="white" )
+ points( pref, 1, pch=1 , col=scol[cls] , lwd=2 )
+ if( sx==1 )
+ {
+ axis(side=2)
+ mtext( "Relative SMR Relative mortality",
+ line=2.5, side=2, outer=FALSE, las=0, cex=0.7 )
+ }
+ box()
+ mtext( c("e","f")[sx], side=3, adj=0.01, line=0.2 )
+ }
}
>
> f1n(cls=1,aod=c(4,6,8,10))
> pdf( "T2Fig1n.pdf", height=9, width=7 )
> f1n( cls=1, aod=c(4,6,8,10) )
> dev.off()
```
All-cause mortality by complication status

pdf

> win.metafile( "T2Fig1n.emf", height=9, width=7 )
> fin( cls=1, aod=c(4,6,8,10) )
> dev.off()

pdf

2
Figure 3.9: Overall mortality rates (a,b), SMR (c,d) and changes in these (e,f). Patients diagnosed with DM in ages 40, 50, 60 and 70. Mortality rates are full lines, SMRs are with broken lines.