Mortality among Type 1 patients at Steno Diabetes Center

SDC
http://BendixCarstensen.com/SDC/EPJmort
Friday 21st June, 2013
Version 2.1

Compiled Friday 21st June, 2013, 18:37
from: C:/Bendix/Steno/MaEJ/EPJ-dod/MortT1.tex

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

Contents

Chapter 1

Data preparation

1.1 Introduction

This report concerns the mortality among type 1 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: C:\Bendix\Steno\MaEJ\EPJ-dod\data

NOTE: AUTOEXEC processing completed.

1          options nofmterr ;
2          libname maej "p:\MAEJ\SAS data\SDC mortality" ;
NOTE: Libref MAEJ was successfully assigned as follows:
      Engine:        V9
      Physical Name: p:\MAEJ\SAS data\SDC mortality
3
```

```
4          title1 "Base dataset - merge of EPJ, NPR and CoDR" ;
5          proc contents  data=maej.compl ;
6          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.

7
8          * Identify fishy records ;
9          data oops nodm notp late compl ;
10           set maej.compl ;
11           if doBth gt doDM gt .z then do ;
12              put "This was changed from: " doBth= ddmmyy10.
13                                           doDM= ddmmyy10. ;
14              doDM = doBth + 90 ;
15              put "                   to: " doBth= ddmmyy10.
16                                           doDM= ddmmyy10. ;
17              end ;
18           if doBth lt "01JAN1900"d then do ;
19              put "This was changed from: " doBth= ddmmyy10.
20                                           doDM= ddmmyy10. ;
21              doBth = mdy( month(doBth), day(doBth), year(doBth)+100 ) ;
22              put "                   to: " doBth= ddmmyy10.
23                                           doDM= ddmmyy10. ;
24              end ;
25           if ( nmiss(doCVD,CVD) eq 1 or
26                nmiss(doDR ,DR ) eq 1 or
27                nmiss(doNef,Nef) eq 1 or
28                nmiss(doNeu,Neu) eq 1 or
29                ( doDth gt .z and CoDth eq "" ) or
30                ( doDth le .z and CoDth ne "" ) or
31                nmiss(sex,doBth,entry,exit) gt 0 or
32                doBth gt entry or
33                entry gt exit )          then output oops ;
34           else if doDM le .z            then output nodm ;
35           else if DMtype eq ""          then output notp ;
36           else if entry gt "31DEC2010"d then output late ;
37           else                               output compl ;
38         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=29/01/1987 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


39
40         title1 "oops" ;
41         proc print  data=oops ;
42           var sex dmtype dobth dodm exit entry doCVD CVD doDR DR doNef Nef doNeu Neu doDth CoDth ;
43           format dobth dodm exit entry doCVD doDR doNef doNeu doDth ddmmyy8. ;
44         run ;

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


45
46         /*
47         title1 "nodm" ;
48         proc print  data=nodm ;
49           var sex dmtype dobth dodm exit entry doCVD CVD doDR DR doNef Nef doNeu Neu doDth CoDth ;
50           format dobth dodm exit entry doCVD doDR doNef doNeu doDth ddmmyy8. ;
51         run ;
52
53         title1 "notp" ;
54         proc print  data=notp ;
55           var sex dmtype dobth dodm exit entry doCVD CVD doDR DR doNef Nef doNeu Neu doDth CoDth ;
56           format dobth dodm exit entry doCVD doDR doNef doNeu doDth ddmmyy8. ;
57         run ;
```

```
58         */
59
60         options validvarname=V6 ;
61         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
62         proc copy in=work out=xptout memtype=data;
63           select compl ;
64         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

Base dataset - merge of EPJ, NPR and CoDR                                10:49 Tuesday, January 29, 2013   1

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             wlatin1  Western (Windows)


                  Engine/Host Dependent Information

Data Set Page Size          12288
Number of Data Set Pages    121
First Data Page             1
Max Obs per Page            95
Obs in First Data Page      76
Number of Data Set Repairs  0
Filename                    p:\MAEJ\SAS data\SDC mortality\compl.sas7bdat
Release Created             9.0202M3
Host Created                W32_VSPRO


Alphabetic List of Variables and Attributes

 #    Variable    Type    Len

10    CVD         Num       8
 8    CoDth       Char      9
 2    DMtype      Char      2
12    DR          Num       8
 7    Entry       Num       8
14    Nef         Num       8
16    Neu         Num       8
 4    doBTH       Num       8
 9    doCVD       Num       8
 5    doDM        Num       8
11    doDR        Num       8
 1    doDTH       Num       8
13    doNef       Num       8
15    doNeu       Num       8
 6    exit        Num       8
 3    sex         Num       8

oops                                                                     10:49 Tuesday, January 29, 2013   2

                                                                                                        Co
Obs sex DMtype    doBTH     doDM     exit     Entry    doCVD CVD     doDR DR    doNef Nef    doNeu Neu    doDTH Dth

 1   1    T2   01/06/47 01/01/08 07/06/12 29/06/10 29/06/10  1         .  .        .  1         .  .         .
 2   1    T2   11/12/71 01/01/10 07/06/12 28/07/10         .  1         .  .        .  .         .  .         .
```

1.2.1 Reading with R

We first load the package needed to read the data:

```
> options( width=100 )
> library( foreign )
> library( Epi )
> library( splines )
> print( sessionInfo(), l=F )


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_1.1.50     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 T1 patients:

```
> epj <- read.xport( "./data/compl.xpt" )
> ( names(epj) <- tolower( names(epj) ) )


 [1] "dodth"  "dmtype" "sex"    "dobth"  "dodm"   "exit"   "entry"  "codth"  "docvd"  "cvd"
[11] "dodr"   "dr"     "donef"  "nef"    "doneu"  "neu"


> table( epj[,2], exclude=NULL )


  T1   T2 <NA>
4855 5876    0


> epj <- subset( epj, dmtype=="T1", select=-2 )
> str( epj )


'data.frame':        4855 obs. of  15 variables:
 $ dodth: num  NA NA NA NA NA ...
 $ sex  : num  1 2 1 2 2 2 1 2 1 2 ...
 $ dobth: num  -10227 -7670 -6574 -3652 -2191 ...
 $ dodm : num  8036 13515 6575 15706 1827 ...
 $ exit : num  19151 19151 19151 19151 19151 ...
 $ entry: num  15063 15063 15063 18324 15063 ...
 $ codth: Factor w/ 10 levels "","accidents",..: 1 1 1 1 1 5 1 1 1 1 ...
 $ docvd: num  14993 NA NA 17512 NA ...
 $ cvd  : num  1 NA NA 1 NA 1 1 NA NA NA ...
 $ dodr : num  15885 13723 12329 NA 12329 ...
 $ dr   : num  1 1 1 NA 1 1 1 NA NA NA ...
 $ donef: num  NA NA NA 16215 NA ...
 $ nef  : num  NA NA NA 1 NA 1 NA NA NA NA ...
 $ doneu: num  15885 NA NA NA 12329 ...
 $ neu  : num  1 NA NA NA 1 NA NA NA NA NA ...
```

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:

```
> 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"
```

```
> 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 )
```

```
     dodth sex dobth    dodm    exit   entry codth   docvd cvd    dodr dr   donef nef   doneu neu
8       NA   M  1932 1982.54 2012.43 2001.24       2001.05   1 2003.49  1      NA  NA 2003.49   1
13      NA   F  1939 1997.33 2012.43 2001.24            NA  NA 1997.57  1      NA  NA      NA  NA
19      NA   M  1942 1978.18 2012.43 2001.24            NA  NA 1993.75  1      NA  NA      NA  NA
30      NA   F  1950 2003.92 2012.43 2010.17       2007.95   1      NA NA 2004.39   1      NA  NA
38      NA   F  1954 1965.76 2012.43 2001.24            NA  NA 1993.75  1      NA  NA 1993.75   1
39 2003.46   F  1955 1971.78 2003.46 2001.24   CVD 1998.32   1 1993.75  1 1993.75   1      NA  NA
```

```
> options(digits=8)
```

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:

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

```
          Entry 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010  Sum
sex Death
M   2001          22    0    0    0    0    0    0    0    0    0   22
    2002          56    0    0    0    0    0    0    0    0    0   56
    2003          32    0    1    0    0    0    0    0    0    0   33
    2004          48    0    1    1    0    0    0    0    0    0   50
    2005          48    1    0    1    1    0    0    0    0    0   51
    2006          52    1    0    1    0    0    0    0    0    0   54
    2007          44    0    2    0    1    1    0    0    0    0   48
    2008          34    0    0    1    0    1    1    0    0    0   37
    2009          29    0    0    0    0    0    2    0    0    0   31
    2010          35    1    2    0    1    2    0    0    1    0   42
    NA          1486   70   60   64   77   91  109  105   87   63 2212
    Sum         1886   73   66   68   80   95  112  105   88   63 2636
F   2001          12    0    0    0    0    0    0    0    0    0   12
    2002          29    0    0    0    0    0    0    0    0    0   29
    2003          36    0    0    0    0    0    0    0    0    0   36
    2004          27    0    1    0    0    0    0    0    0    0   28
    2005          28    0    0    0    0    0    0    0    0    0   28
    2006          21    0    1    0    0    0    0    0    0    0   22
    2007          34    0    1    1    0    0    0    0    0    0   36
    2008          37    0    0    0    0    0    0    1    0    0   38
    2009          21    1    0    0    0    0    1    1    1    0   25
    2010          28    0    0    0    0    0    0    0    1    0   29
    NA          1347   47   51   75   56   50   77   99   62   72 1936
```

```
    Sum          1620   48   54   76   56   50   78  101   64   72 2219
Sum 2001           34    0    0    0    0    0    0    0    0    0   34
    2002           85    0    0    0    0    0    0    0    0    0   85
    2003           68    0    1    0    0    0    0    0    0    0   69
    2004           75    0    2    1    0    0    0    0    0    0   78
    2005           76    1    0    1    1    0    0    0    0    0   79
    2006           73    1    1    1    0    0    0    0    0    0   76
    2007           78    0    3    1    1    1    0    0    0    0   84
    2008           71    0    0    1    0    1    1    1    0    0   75
    2009           50    1    0    0    0    0    3    1    1    0   56
    2010           63    1    2    0    1    2    0    0    2    0   71
    NA           2833  117  111  139  133  141  186  204  149  135 4148
    Sum          3506  121  120  144  136  145  190  206  152  135 4855
```

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:

```
> 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 2010-12-24 So we just check if all persons enter and exits correctly:

```
> all.exit <- cal.yr( "2010-12-31" )
> 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             424  283  707
NA              2212 1936 4148
Sum             2636 2219 4855
```

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

```
> tt <- with( epj, addmargins( table(DN=round(donef,3),
+                                    Dead=!is.na(dodth),
+                                      exclude=NULL) ) )
> tt[tt[,"Sum"]>5,]
```

```
          Dead
DN          FALSE TRUE <NA>  Sum
  1993.755   434  206    0  640
  1994.001     4    2    0    6
  1999.001     9    0    0    9
  2005.747    66    4    0   70
  2005.749     9    0    0    9
  2006.001     8    0    0    8
  <NA>      3168  351    0 3519
  Sum       4148  707    0 4855
```

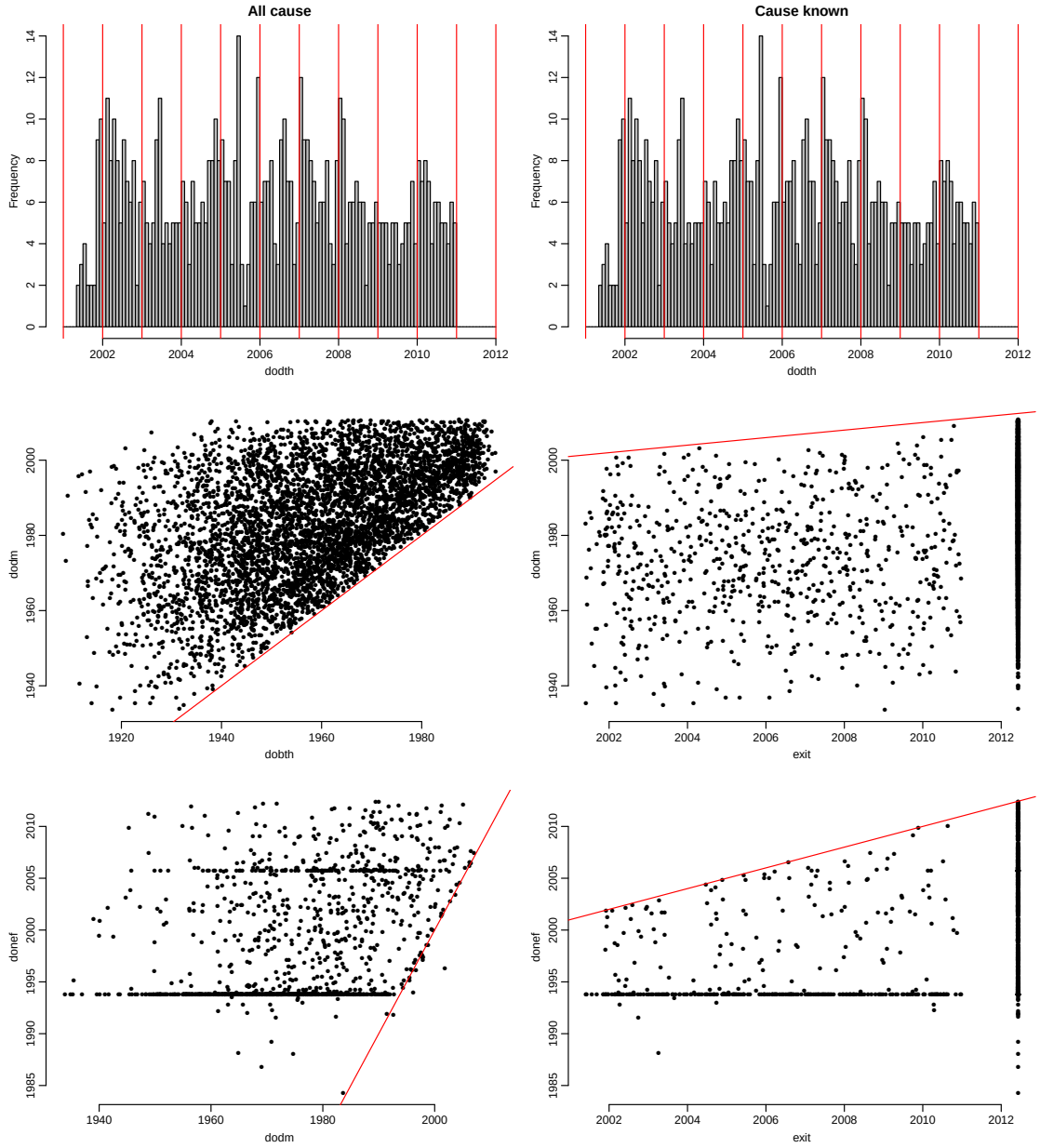

Figure 1.1: *Histogram of dates of death for all known deaths and for deaths where a cause is known. The bottom 4 panels have plots of the date of diagnosis resp. complications versus date of birth and date of exit, with a red line indicating the identity (so all points should preferably be on the same 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 uncertainly. 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:

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

```
codth
          accidents  acute DM    Cancer       CVD        GI Infection    kidney      Lung     other
     4148        22        19        98       223        34        56        18        35       202
```

```
> # 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$codth <- Relevel( epj$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 ) )
```

```
            CoD
codth        Alive  CVD Cancer Other
              4148    0      0     0
  CVD            0  223      0     0
  Cancer         0    0     98     0
  Acute DM       0    0      0    19
  Kidney         0    0      0    18
  Lung           0    0      0    35
  GI             0    0      0    34
  Infection      0    0      0    56
  Accidents      0    0      0    22
  Other          0    0      0   202
```

```
> with( epj, ftable( addmargins( table( sex,
+                                       doDTH=floor(dodth), CoD,
+                                       useNA="ifany" ),
+                                margin=2:3), row.vars=2 ) )
```

```
      sex     M                              F
      CoD Alive  CVD Cancer Other  Sum Alive  CVD Cancer Other  Sum
doDTH
2001          0   10      2    10   22     0    7      0     5   12
2002          0   24      5    27   56     0   11      3    15   29
2003          0    8      7    18   33     0   14      4    18   36
2004          0   17      7    26   50     0    9      4    15   28
2005          0   19      3    29   51     0    7      9    12   28
```

```
2006          0   15      9    30   54     0    9      4     9   22
2007          0    9      5    34   48     0    7      6    23   36
2008          0    8      5    24   37     0   10      7    21   38
2009          0   15      2    14   31     0    4      1    20   25
2010          0   10      8    24   42     0   10      7    12   29
NA         2212    0      0     0 2212  1936    0      0     0 1936
Sum        2212  135     53   236 2636  1936   88     45   150 2219
```

```
> with( epj, ftable( addmargins( table( sex,
+                                       doDTH=floor(dodth), Dead=!is.na(dodth),
+                                       useNA="ifany" ) ),
+                    row.vars=2 ) )
```

```
      sex      M                F             Sum
      Dead FALSE TRUE  Sum FALSE TRUE  Sum FALSE TRUE  Sum
doDTH
2001          0   22   22     0   12   12     0   34   34
2002          0   56   56     0   29   29     0   85   85
2003          0   33   33     0   36   36     0   69   69
2004          0   50   50     0   28   28     0   78   78
2005          0   51   51     0   28   28     0   79   79
2006          0   54   54     0   22   22     0   76   76
2007          0   48   48     0   36   36     0   84   84
2008          0   37   37     0   38   38     0   75   75
2009          0   31   31     0   25   25     0   56   56
2010          0   42   42     0   29   29     0   71   71
NA         2212    0 2212  1936    0 1936  4148    0 4148
Sum        2212  424 2636  1936  283 2219  4148  707 4855
```

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):

```
> 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.

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

$M

Transitions:
     To
From    Alive Dead  Records:  Events: Risk time:  Persons:
  Alive  2212  424       2636      424   19516.34      2636

$F

Transitions:
     To
From    Alive Dead  Records:  Events: Risk time:  Persons:
  Alive  1936  283       2219      283   16930.43      2219
```

1.3.1 Raw mortality by calendar year

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

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


Transitions:
     To
From     Alive Dead  Records:  Events: Risk time:  Persons:
  Alive 149486  707    150193      707   36446.78      4855
```

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

```
> 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 )

                    sex     M                 F
                            D     Y  rate     D     Y  rate
I(floor(per * 4)/4)
2001                      0.0  17.8   0.0   0.0  15.3   0.0
2001.25                   2.0 458.7   0.4   3.0 393.7   0.8
2001.5                    4.0 463.6   0.9   4.0 396.1   1.0
2001.75                  16.0 466.6   3.4   5.0 399.7   1.3
2002                     15.0 466.2   3.2   9.0 402.3   2.2
2002.25                  15.0 467.6   3.2   8.0 402.5   2.0
2002.5                   13.0 469.1   2.8   9.0 403.1   2.2
2002.75                  13.0 469.9   2.8   3.0 405.6   0.7
2003                      7.0 471.2   1.5   9.0 407.1   2.2
2003.25                  13.0 474.3   2.7  12.0 408.6   2.9
2003.5                    6.0 475.7   1.3   7.0 409.1   1.7
2003.75                   7.0 478.2   1.5   8.0 410.8   1.9
2004                      9.0 479.3   1.9   7.0 412.4   1.7
2004.25                  12.0 481.8   2.5   5.0 414.8   1.2
2004.5                   15.0 481.0   3.1   4.0 417.6   1.0
2004.75                  14.0 482.6   2.9  12.0 422.2   2.8
2005                     13.0 483.0   2.7  10.0 423.4   2.4
2005.25                  16.0 483.1   3.3   9.0 423.6   2.1
2005.5                    5.0 483.2   1.0   2.0 425.1   0.5
2005.75                  17.0 489.7   3.5   7.0 428.2   1.6
2006                     16.0 492.0   3.3   4.0 430.9   0.9
2006.25                   7.0 495.4   1.4   8.0 432.7   1.8
2006.5                   20.0 497.0   4.0   6.0 434.0   1.4
2006.75                  11.0 499.4   2.2   4.0 436.0   0.9
2007                     18.0 502.4   3.6  12.0 438.4   2.7
2007.25                  12.0 505.6   2.4   8.0 440.4   1.8
2007.5                    9.0 509.4   1.8  11.0 442.4   2.5
2007.75                  10.0 514.2   1.9   5.0 446.3   1.1
2008                     10.0 518.8   1.9  15.0 449.8   3.3
2008.25                   7.0 523.8   1.3  12.0 453.8   2.6
2008.5                    7.0 528.9   1.3   7.0 457.2   1.5
2008.75                  12.0 533.1   2.3   4.0 461.5   0.9
2009                     10.0 535.5   1.9   5.0 465.4   1.1
2009.25                   7.0 538.2   1.3   6.0 469.4   1.3
2009.5                    5.0 542.3   0.9   7.0 470.8   1.5
2009.75                   9.0 545.7   1.6   7.0 472.3   1.5
```

```
2010                  13.0 548.3   2.4  10.0 474.4   2.1
2010.25               10.0 549.6   1.8   7.0 477.5   1.5
2010.5                10.0 549.8   1.8   6.0 480.1   1.2
2010.75                9.0 544.4   1.7   6.0 476.0   1.3
```

```
> 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()
```

A quick look at the tables or figure 1.2 shows that data seem incomplete prior to 2002, possibly not for the last 3 months of 2001, but to make sure that data are valid, we define entry to be at the start of 2002, and redefine the Lexis object:

```
> 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.
```

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

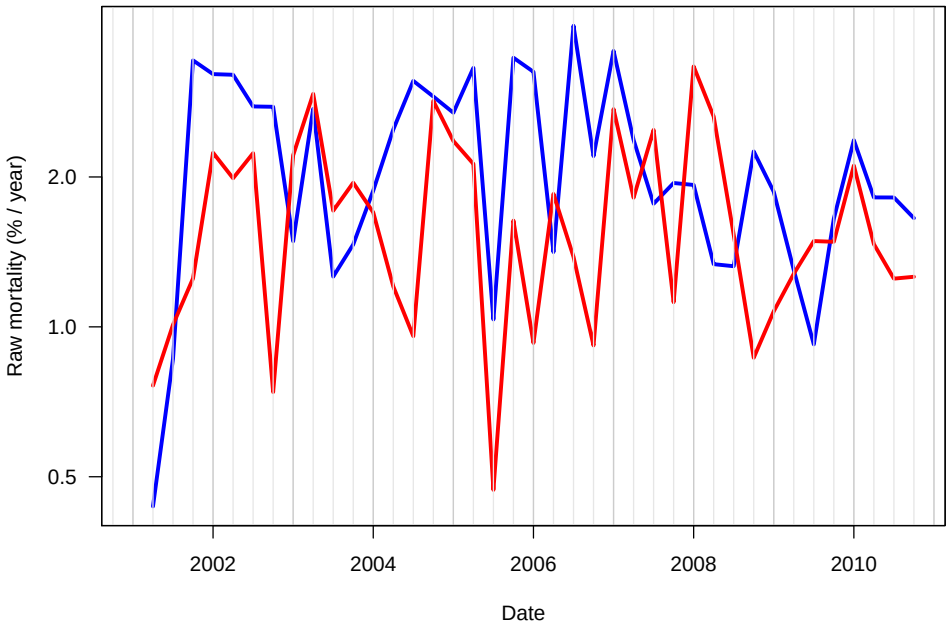

Figure 1.2: *Raw mortality rates for T1 patients by 3-month periods; there is something missing prior to 2002.*

```
$M

Transitions:
     To
From    Alive Dead  Records:  Events: Risk time:  Persons:
  Alive  2212  402      2614      402   18109.62      2614

$F

Transitions:
     To
From    Alive Dead  Records:  Events: Risk time:  Persons:
  Alive  1936  271      2207      271   15725.69      2207
```

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

```
      age              Ddur              per            lex.dur           lex.Cst      lex.Xst
 Min.   : 9.7399   Min.   : 0.0000   Min.   :2002.0   Min.   :0.0013689   Alive:4821   Alive:4148
 1st Qu.:31.6208   1st Qu.: 7.5415   1st Qu.:2002.0   1st Qu.:4.9295003   Dead :   0   Dead : 673
 Median :43.1718   Median :17.8814   Median :2002.0   Median :8.9965777
 Mean   :44.2820   Mean   :19.4668   Mean   :2003.3   Mean   :7.0183189
 3rd Qu.:56.0452   3rd Qu.:29.2667   3rd Qu.:2003.2   3rd Qu.:8.9965777
 Max.   :93.6030   Max.   :68.4094   Max.   :2011.0   Max.   :8.9965777

     lex.id          dodth        sex          dobth             dodm             exit
 Min.   :   1   Min.   :2002.0   M:2614   Min.   :1908.4   Min.   :1933.6   Min.   :2002.0
 1st Qu.:1206   1st Qu.:2004.2   F:2207   1st Qu.:1946.4   1st Qu.:1973.2   1st Qu.:2012.4
 Median :2411   Median :2006.3            Median :1959.6   Median :1985.1   Median :2012.4
 Mean   :2411   Mean   :2006.3            Mean   :1959.0   Mean   :1983.9   Mean   :2011.6
 3rd Qu.:3616   3rd Qu.:2008.4            3rd Qu.:1972.0   3rd Qu.:1996.4   3rd Qu.:2012.4
 Max.   :4821   Max.   :2011.0            Max.   :1994.8   Max.   :2010.9   Max.   :2012.4
                NA's   :4148
     entry               codth          docvd             cvd            dodr              dr
 Min.   :2002.0            :4148   Min.   :1986.7   Min.   :1      Min.   :1983.2   Min.   :1
 1st Qu.:2002.0   CVD      : 206   1st Qu.:1993.8   1st Qu.:1      1st Qu.:1993.8   1st Qu.:1
 Median :2002.0   Other    : 195   Median :2001.5   Median :1      Median :1993.8   Median :1
 Mean   :2003.3   Cancer   :  96   Mean   :2001.1   Mean   :1      Mean   :1997.9   Mean   :1
 3rd Qu.:2003.2   Infection:  54   3rd Qu.:2005.9   3rd Qu.:1      3rd Qu.:2002.4   3rd Qu.:1
 Max.   :2011.0   GI       :  33   Max.   :2012.4   Max.   :1      Max.   :2012.4   Max.   :1
                  (Other)  :  89   NA's   :2188     NA's   :2188   NA's   :1526     NA's   :1526
     donef             nef           doneu             neu            CoD
 Min.   :1984.3   Min.   :1      Min.   :1987.1   Min.   :1      Alive :4148
 1st Qu.:1993.8   1st Qu.:1      1st Qu.:1993.8   1st Qu.:1      CVD   : 206
 Median :1994.0   Median :1      Median :1994.0   Median :1      Cancer:  96
 Mean   :1998.3   Mean   :1      Mean   :1998.0   Mean   :1      Other : 371
 3rd Qu.:2003.4   3rd Qu.:1      3rd Qu.:2002.4   3rd Qu.:1
 Max.   :2012.3   Max.   :1      Max.   :2012.4   Max.   :1
 NA's   :3500     NA's   :3500   NA's   :3078     NA's   :3078
```

```
> save( L1, file="./data/T1L1.Rda" )
```

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(0,90)
> 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$counts[1], font=2, adj=-0.1, col="blue" )
> zz <-
+ with( subset(L1, sex=="F" ),
+       hist( Ddur, breaks=0:100, col="red", border="red",
+             main="", ylim=yl, xlab="", ylab="" ) )
> text( 1.5, yl[2], zz$counts[1], font=2, adj=-0.1, col="red" )
> mtext( c("Age at entry","DM duration at entry"), side=3, line=0,
+        at=c(1,3)/4, outer=TRUE )
```

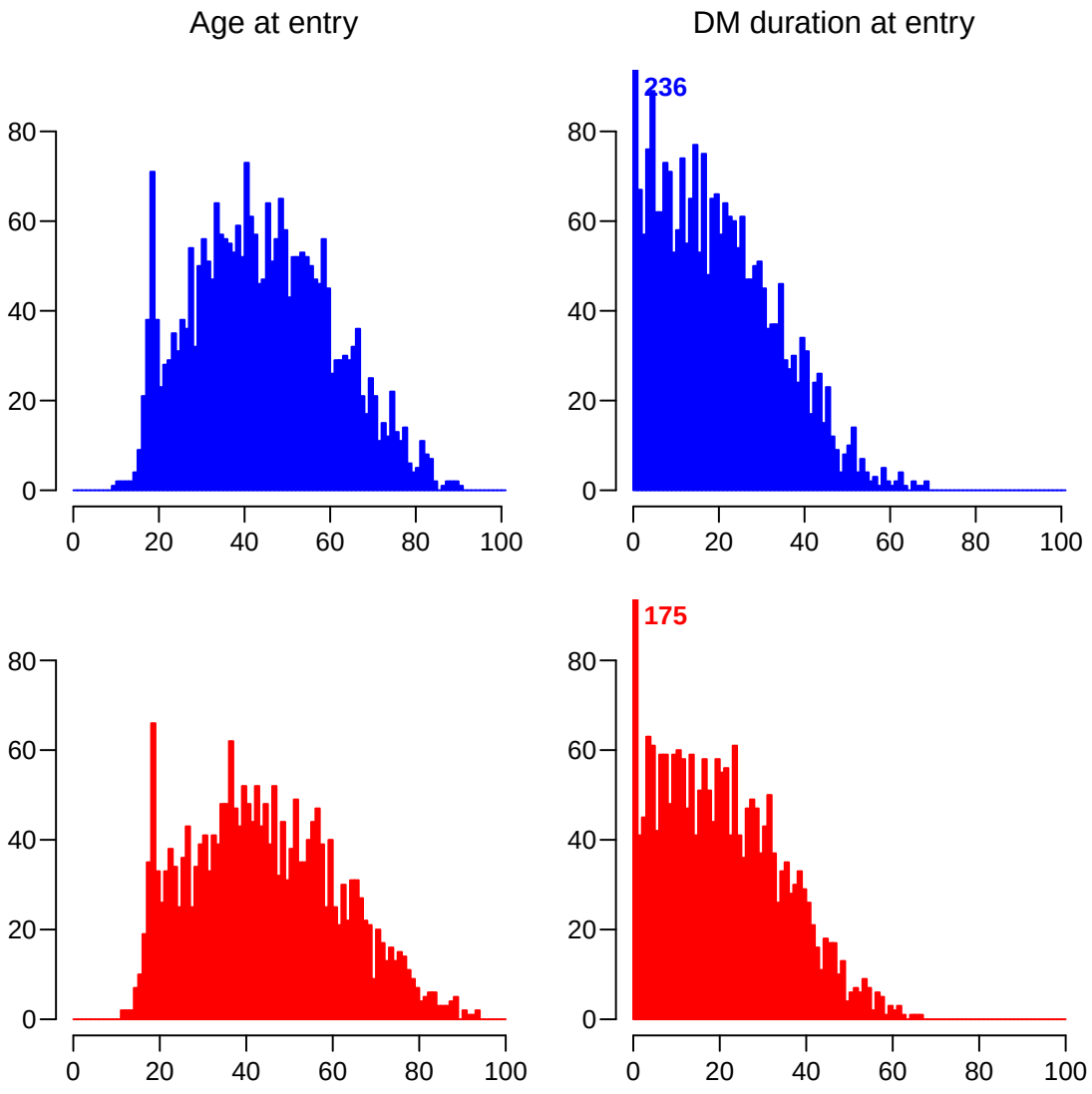

Figure 1.3: *Entry age and duration for T1 patients at Steno.*

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   17      2    15
  2002     0   35      8    42
  2003     0   22     11    36
  2004     0   26     11    41
  2005     0   26     12    41
  2006     0   24     13    39
  2007     0   16     11    57
  2008     0   18     12    45
  2009     0   19      3    34
  2010     0   20     15    36
  2012  4148    0      0     0
```

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

```
       entryOK
deathOK TRUE
   TRUE  707
   <NA> 4148
```

```
> 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] 4821
```

```
> summary( C1 )
```

```
Transitions:
     To
From    Alive CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive  4148 206     96   371      4821      673   33835.32      4821
```

```
> summary( C1, by="sex" )
```

```
$M

Transitions:
     To
From    Alive CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive  2212 125     51   226      2614      402   18109.62      2614

$F

Transitions:
     To
From    Alive CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive  1936  81     45   145      2207      271   15725.69      2207


> save( C1, file="./data/T1CoD.Rda" )
```

1.5 Base tables

Once we have groomed the `L1` dataset we can start by making the baseline table (table 1): The state of DN is defined as presence of complications 180 days after entry:

```
> source("c:/stat/r/bxc/library.sources/useful/R/NArray.r")
> NArray


function (x)
{
    if (!is.list(x))
        stop("Argument must be a (named) list.")
    array(NA, dimnames = x, dim = sapply(x, length))
}


> tab1 <- NArray(
+         list( sex = levels(L1$sex),
+                     c("N","Age","Age-IQR","DM dur","DMdur-IQR",
+                       "Nephropathy",
+                       "Neuropathy",
+                       "Retinopathy",
+                       "CVD",
+                       "Deaths"),
+                     c("N","%/sd") ) )
> str( tab1 )


 logi [1:2, 1:10, 1:2] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 3
  ..$ sex: chr [1:2] "M" "F"
  ..$    : chr [1:10] "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["F","N","%/sd"] <- tab1["F","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[,"DM dur","%/sd"] <- with( L1, tapply(entry+0.5-dodm,sex,sd) )
> 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[,"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[,6:10,"%/sd"] <- tab1[,6:10,"N"] / tab1[,"N",rep(1,5)] * 100
> round( ftable(tab1,col.vars=c(1,3)), 1 )
```

```
            sex      M             F
                     N   %/sd      N   %/sd

N               2614.0   54.2 2207.0   45.8
Age               43.8   16.1   43.4   17.0
Age-IQR           32.5   56.2   31.6   57.0
DM dur            17.6   14.1   19.0   14.2
DMdur-IQR          7.6   29.1    8.6   30.6
Nephropathy      612.0   23.4  422.0   19.1
Neuropathy       843.0   32.2  578.0   26.2
Retinopathy     1469.0   56.2 1240.0   56.2
CVD              956.0   36.6  707.0   32.0
Deaths           402.0   15.4  271.0   12.3
```

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

```
            sex    M         F
                   N %/sd    N %/sd

N               2614   54 2207   46
Age               44   16   43   17
Age-IQR           33   56   32   57
DM dur            18   14   19   14
DMdur-IQR          8   29    9   31
Nephropathy      612   23  422   19
Neuropathy       843   32  578   26
Retinopathy     1469   56 1240   56
CVD              956   37  707   32
Deaths           402   15  271   12
```

Chapter 2

Mortality by cause of death

```
> load( file="./data/T1CoD.Rda" )
```

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

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

```
$M

Transitions:
     To
From    Alive CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive 72656 125     51   226     73058      402   18109.62      2614

$F

Transitions:
     To
From    Alive CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive 63090  81     45   145     63361      271   15725.69      2207
```

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

```
> 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:

```
> xl <- c(0,80)
> yl <- c(0,100)
> ypi <- 16
> diff(xl)/ypi
```

```
[1] 5
```

```
> diff(yl)/ypi
```

```
[1] 6.25
```

```
> 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(0,80), 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$lex.Xst] )
> rect( 65,0,80,20, col="white", border="lightgray" )
> text( rep(78,5), 1:4*3.5, c(levels(S1$lex.Xst)[2:4],"C.o.D."),
+       col=c(rainbow(3),"gray"), adj=1, cex=0.9, font=2 )
> box()
```

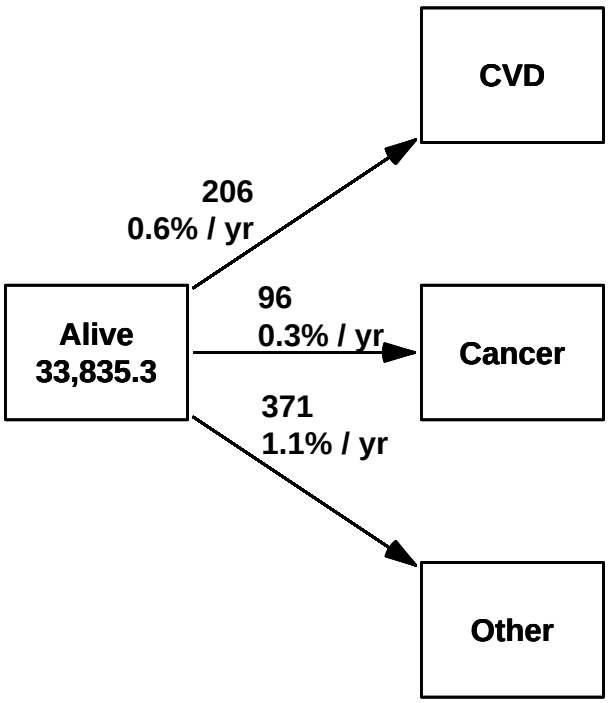

Figure 2.1: *Person-years, deaths and mortality rates by cause of death in the SDC T1 patient population.*

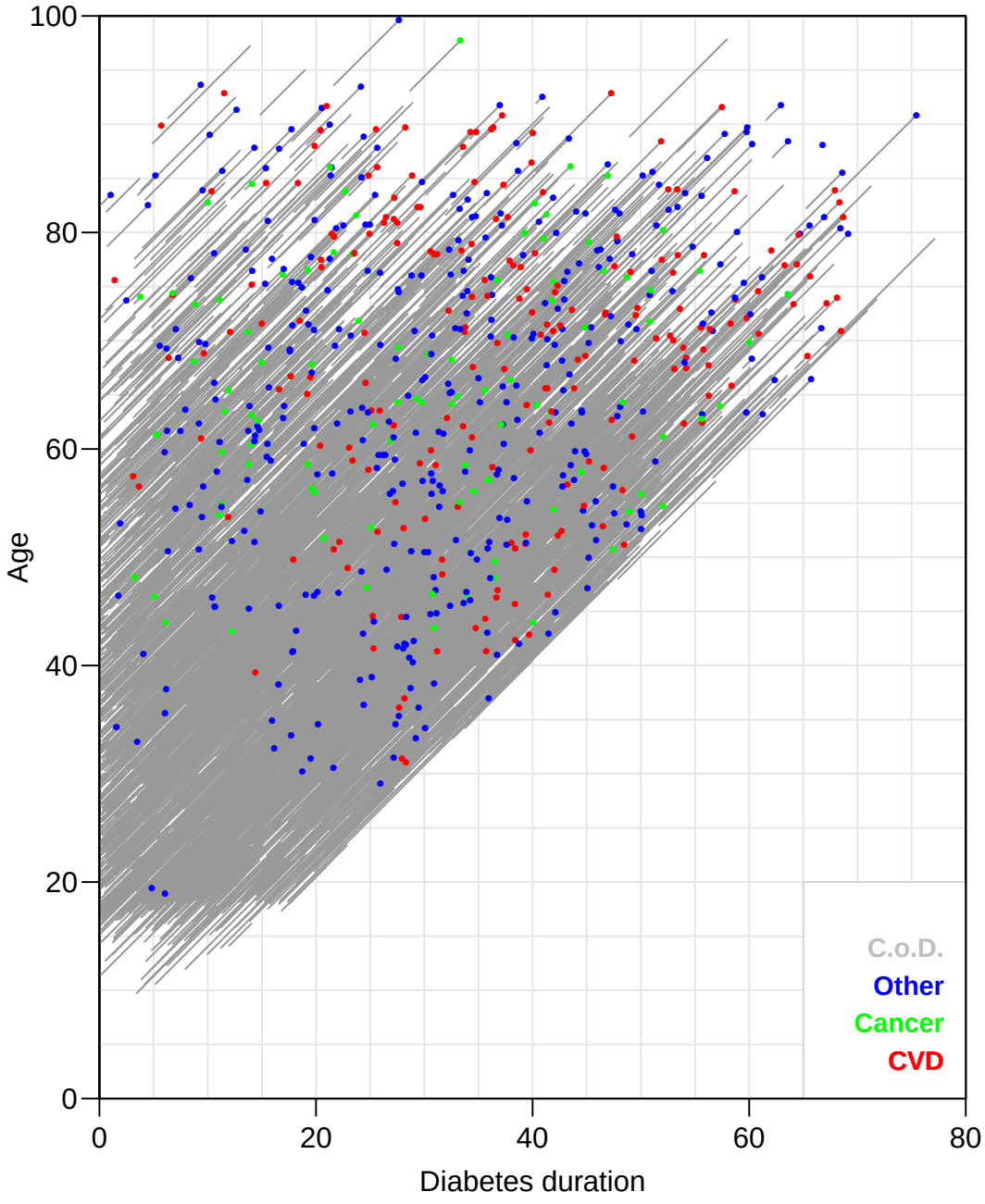

Figure 2.2: *Distribution of follow-up and deaths (by cause) for T1. Although not visible directly, no person in this plot has a life-line (total follow-up) of more than 7 years, since the earliest entry is 1.1.2002, and the latest exit is 31.12.2008.*

2.1 Statistical analysis

2.1.1 Setup

For practical modeling proposes we need a simple way of specifying natural splines, so we set up a wrapper for the function `ns` that automatically takes the smallest and largest knots as boundary knots without further ado:

```
> source("c:/stat/r/bxc/library.sources/useful/r/NS.R")
> Ns


function (x, df = NULL, knots = NULL, intercept = FALSE, Boundary.knots = NULL)
{
    if (is.null(Boundary.knots)) {
        if (!is.null(knots)) {
            knots <- sort(unique(knots))
            ok <- c(1, length(knots))
            Boundary.knots <- knots[ok]
            knots <- knots[-ok]
        }
    }
    ns(x, df = df, knots = knots, intercept = intercept, Boundary.knots = Boundary.knots)
}
```

Then we set up the modeling parameters for the age, period and duration effects:

```
> 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%
45.434086 58.518823 67.707050 75.452977 84.333470


> 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%
2002.7461 2006.3217 2010.0304


> 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  25.793025  39.946270


> 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:

```
> 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:

```
> 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
  34.45    3.71   38.93
```

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

```
           what P(lin) RR / year    lo    hi P(null)
cod    sex
CVD    M         0.944     0.899 0.840 0.963   0.002
       F         0.336     0.904 0.831 0.984   0.019
Cancer M         0.785     0.952 0.856 1.058   0.358
       F         0.185     1.007 0.900 1.127   0.903
Other  M         0.069     0.958 0.912 1.008   0.097
       F         0.732     0.982 0.922 1.045   0.562
```

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

```
           what RR / year    lo    hi
cod    sex
CVD    M            -10.1 -16.0  -3.7
       F             -9.6 -16.9  -1.6
Cancer M             -4.8 -14.4   5.8
       F              0.7 -10.0  12.7
Other  M             -4.2  -8.8   0.8
       F             -1.8  -7.8   4.5
```

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 10%/year, not detectable for cancer, and for other causes only a borderline significant decrease of some 5%/year, but only for men. Thus it seems that the major change in mortality rates among the T1D patients is for CVD mortality.

We can now plot the age-effects:

```
> mlim <- c(1,1000)/30
> rlim <- c(1/3,3)
> par( mfcol=c(2,3), 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(res)[["cod"]] )
+    {
+ matplot( a.pr, a.pr, type="n", log="y", ylim=mlim, yaxt="n",
+          xlab="Age", ylab="", las=1 )
+ if( cd==dimnames(res)[["cod"]][1] ) axis(side=2)
+ abline( v=seq(0,100,5), h=outer(1:9,10^(-2:1)), col=gray(0.8) )
+
+ for( sx in 1:2 )
+ matlines( a.pr, res["Ainc",cd,sx,,],
+           lwd=c(3,1,1), lty=1, col=c("blue","red")[sx] )
+ mtext( cd, line=0.5, side=3, outer=FALSE )
+ if( cd==dimnames(res)[["cod"]][1] )
+     {
+     axis( side=2 )
+     mtext( "Mortality at 1 Jan 2008 (%/year)", line=3, side=2,
+             outer=FALSE, las=0 )
+     }
+
+ matplot( p.pr, p.pr, type="n", log="y", ylim=rlim, las=1,
+          xlab="Date of follow-up", ylab="", yaxt="n" )
+ abline( v=2000+0:11, h=c(1:15/10,1:15,1:15*10), col=gray(0.8) )
+
+ for( sx in 1:2 )
+ matlines( p.pr, res["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( pref ,1, pch=1, lwd=2 )
+ if( cd==dimnames(res)[["cod"]][1] )
+     {
+     axis( side=2 )
+     mtext( "Mortality RR", line=3, side=2, outer=FALSE, las=0 )
+     }
+ }
```

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.

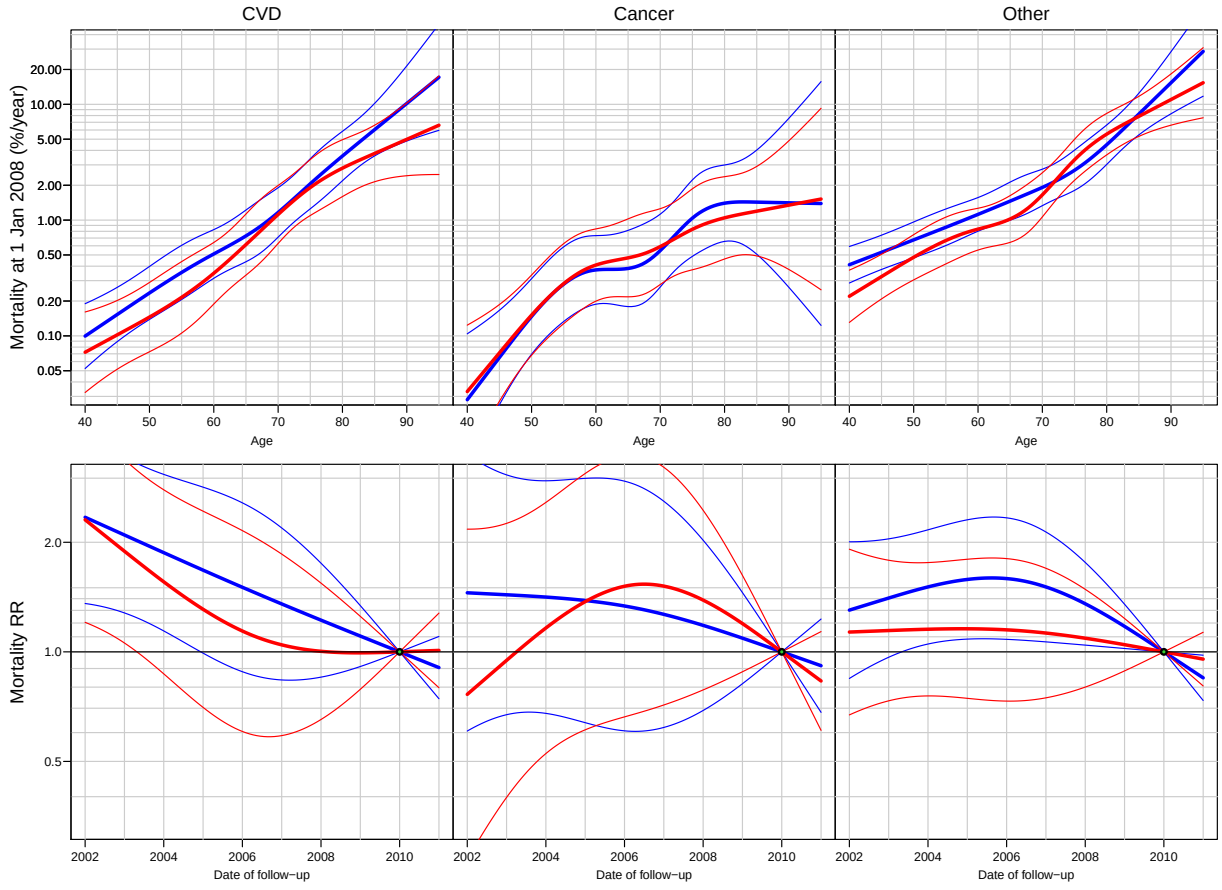

Figure 2.3: *Age-specific mortality and the change of this over calendar time for T1 diabetes patients. Red: F; blue: M, thin lines: 95% c.i.*

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:

```
> 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"
```

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

```
 logi [1:3, 1:2, 1:2, 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 [1:2] "PRR" "DRR"
  ..$ what: chr [1:4] "P(lin)" "RR / year" "lo" "hi"
```

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

```
> system.time(
+ for( cd in dimnames(resx)[["cod"]] )
+ for( sx in dimnames(resx)[["sex"]] )
+ {
+ mapd <- glm( (lex.Xst == cd) ~ Ns( per, knots=p.kn )
+                              + Ns( age, knots=a.kn )
+                              + Ns( Ddur, knots=d.kn ),
+              offset = log(lex.dur/100),
+              family = poisson,
+                data = subset( S1, sex==sx ) )
+ 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" )[2:3,"Pr(>Chi)"],
+                         rbind( ci.exp( mapl, subset="per" ),
+                                ci.exp( madl, subset="Ddur" ) ) )
+ 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
  52.15    0.69   53.80
```

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:

```
> round( ftable( linx, col.vars=3:4 ), 3 )
```

```
           eff     PRR                          DRR
           what P(lin) RR / year    lo    hi P(lin) RR / year    lo    hi
cod    sex
CVD    M         0.912     0.892 0.833 0.955  0.260     1.027 1.015 1.040
       F         0.310     0.895 0.822 0.975  0.051     1.024 1.009 1.039
Cancer M         0.783     0.955 0.859 1.061  0.701     0.986 0.968 1.005
       F         0.187     1.003 0.896 1.123  0.598     1.007 0.988 1.027
Other  M         0.069     0.957 0.910 1.006  0.146     1.004 0.995 1.013
       F         0.745     0.975 0.916 1.039  0.486     1.011 1.000 1.022
```

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

```
           eff        PRR                   DRR
           what RR / year    lo    hi RR / year    lo    hi
cod    sex
CVD    M           -10.8 -16.7  -4.5       2.7   1.5   4.0
       F           -10.5 -17.8  -2.5       2.4   0.9   3.9
Cancer M            -4.5 -14.1   6.1      -1.4  -3.2   0.5
       F             0.3 -10.4  12.3       0.7  -1.2   2.7
Other  M            -4.3  -9.0   0.6       0.4  -0.5   1.3
       F            -2.5  -8.4   3.9       1.1   0.0   2.2
```

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 +2.5%/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 1%/year, but only for women.

```
> mlim <- c(1,1000)/30
> rlim <- c(1/5,5)
> par( mfcol=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", ylab="", 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=c(3,1,1), lty=1, col=c("blue","red")[sx] )
+ mtext( cd, line=0.5, side=3, outer=FALSE )
+ if( cd==dimnames(resx)[["cod"]][1] )
+     {
+     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, las=1, yaxt="n",
+          xlab="Date of follow-up", ylab="" )
+ abline( v=2000+1:15, h=outer(1:9,10^(-2:2)), col=gray(0.8) )
+ for( sx in 1:2 )
+ 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 )
+     }
+  }
```

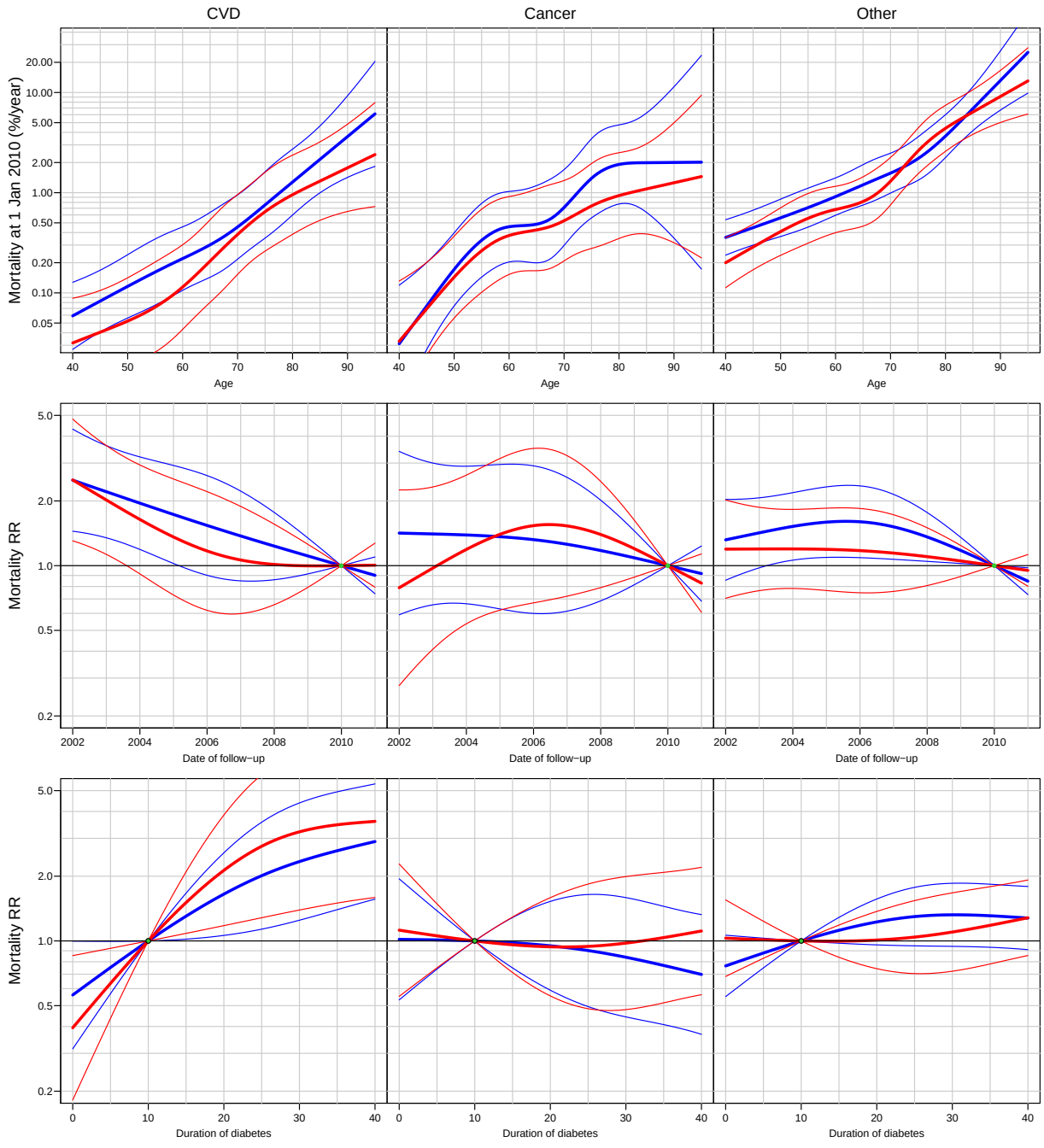

Figure 2.4: *Age-specific mortality and the change of this over calendar time for T1 patients. Red: F; blue: M, thin lines: 95% c.i.*

Chapter 3

Analysis of all-cause mortality by complication status

We start by making a tabular overview of how patients go through the study:

```
> load( file="./data/T1CoD.Rda" )
> with( C1, ftable( addmargins( table( sex,
+                                      eDN=!(donef>entry | is.na(donef)),
+                                      lex.Xst,
+                                      xD=!is.na(donef) ),
+                               margin = c(1,4) ),
+                   col.vars=4:3 ) )
```

```
          xD      FALSE                    TRUE                     Sum
          lex.Xst Alive  CVD Cancer Other Alive  CVD Cancer Other Alive  CVD Cancer Other
sex eDN
M   FALSE          1655   44     34   107   149    8      4    18  1804   52     38   125
    TRUE              0    0      0     0   408   73     13   101   408   73     13   101
F   FALSE          1513   40     32    75   120    3      1    11  1633   43     33    86
    TRUE              0    0      0     0   303   38     12    59   303   38     12    59
Sum FALSE          3168   84     66   182   269   11      5    29  3437   95     71   211
    TRUE              0    0      0     0   711  111     25   160   711  111     25   160
```

3.1 Setup

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

```
> C2 <- cutLexis( C1, cut = L1$donef,
+                 timescale = "per",
+                 new.state = "DN",
+                 new.scale = "DNdur",
+          precursor.states = "Alive" )
> summary( C2, by="sex", scale=100 )
```

```
$M

Transitions:
     To
From    Alive  DN CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive  1668 166  44     34   107      2019      351     132.27      2019
  DN        0 544  81     17   119       761      217      48.83       761
  Sum    1668 710 125     51   226      2780      568     181.10      2614
```

```
$F

Transitions:
     To
From    Alive  DN CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive  1525 123  40     32    75      1795      270     120.51      1795
  DN        0 411  41     13    70       535      124      36.75       535
  Sum    1525 534  81     45   145      2330      394     157.26      2207


> timeScales( C2 )

[1] "age"   "Ddur"  "per"   "DNdur"
```

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:

```
> ypi <- 14
> ll  <- 50
```

We can then set up the plot correctly in a `pdf`-file:

```
> pdf( "./graph/T1mort-Lexis-dur.pdf", height=100/ypi+1, width=80/ypi+1 )
> par( mai=c(3,3,1,1)/4, mgp=c(3,1,0)/1.5, las=1, oma=c(0,0,0,0) )
> clr <- c(gray(c(0.8,0.5)),"Black")
> idt1 <- C1$lex.id
>  nt1 <- length( idt1 )
> sbt1 <- sample( idt1, floor(nt1/1) )
> plsb <- subset( C2, lex.id %in% sbt1 )
> plot.Lexis( C2, time.scale=2:1, grid=1:20*5, lty.grid=1,
+             col=clr[plsb$lex.Cst],lwd=2,
+             xlim=c(0,80),ylim=c(0,100),xaxs="i",yaxs="i",
+             ylab="Age",xlab="Duration of diabetes")
> points( C2, pch=16, cex=0.5,
+         col=c(rep("transparent",2),rainbow(3))[C2$lex.Xst] )
> rect( 65,0,80,20, col="white", border="lightgray" )
> text( rep(78,5), 1:4*3.5, c(levels(C2$lex.Xst)[3:5],"C.o.D."),
+       col=c(rainbow(3),gray(0.4)), adj=1, cex=0.9, font=2 )
> box()
> dev.off()

null device
          1
```

We will also want to see the number of transitions between states:

```
> par(mfrow=c(1,2), mar=c(0,0,0,0) )
> boxes( C2, boxpos=list( x=c(15,15,85,85,85),
+                         y=c(80,20,10,50,90) ), scale.R=100,
+        pos.arr=c(5,8,8,8,8,9,8)/10 )
> C3 <- Relevel( C2, list( Dead=3:5 ), first=FALSE )

      type    old   new
1  lex.Cst  Alive Alive
2  lex.Cst     DN    DN
3  lex.Cst    CVD
4  lex.Cst Cancer
5  lex.Cst  Other
6  lex.Xst  Alive Alive
7  lex.Xst     DN    DN
8  lex.Xst    CVD  Dead
9  lex.Xst Cancer  Dead
10 lex.Xst  Other  Dead
```

```
> boxes( C3, boxpos=list( x=c(15,15,85),
+                         y=c(80,20,50) ), scale.R=100,
+        pos.arr=c(5,8,9)/10 )
```

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.

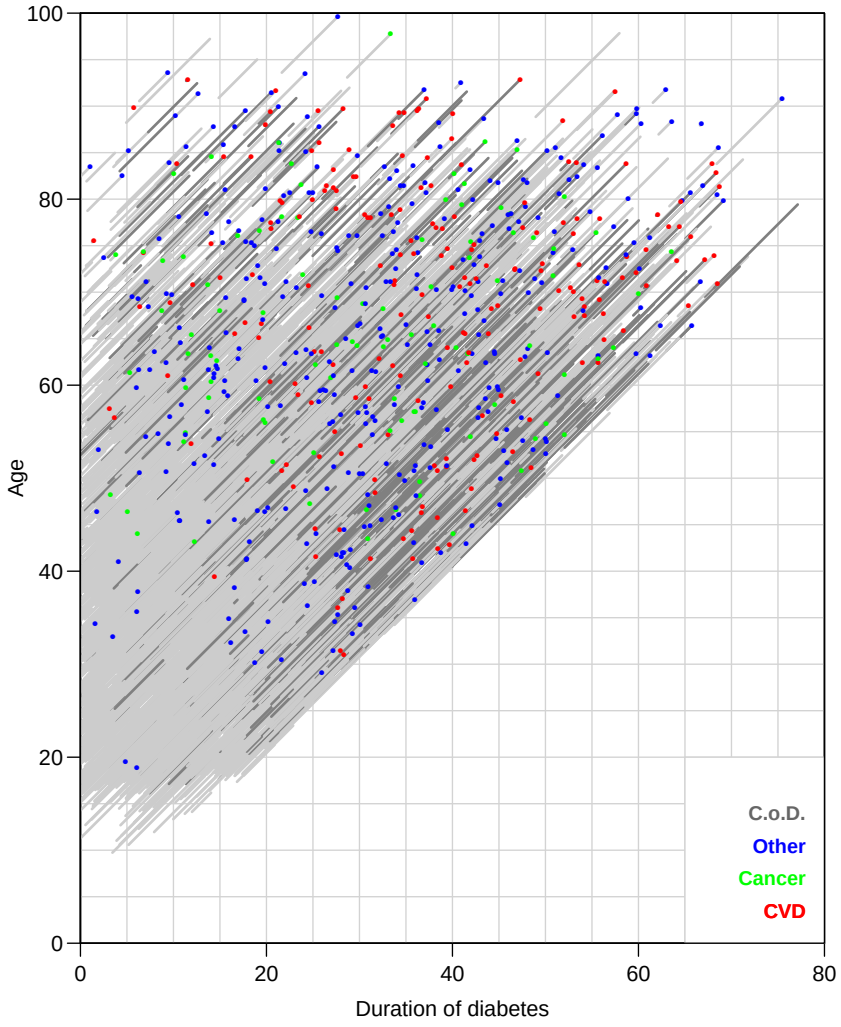

Figure 3.1: *Follow-up for T1 patients at Steno. Follow-up after onset of DN is shown in dark gray color, deaths shown as dots.*

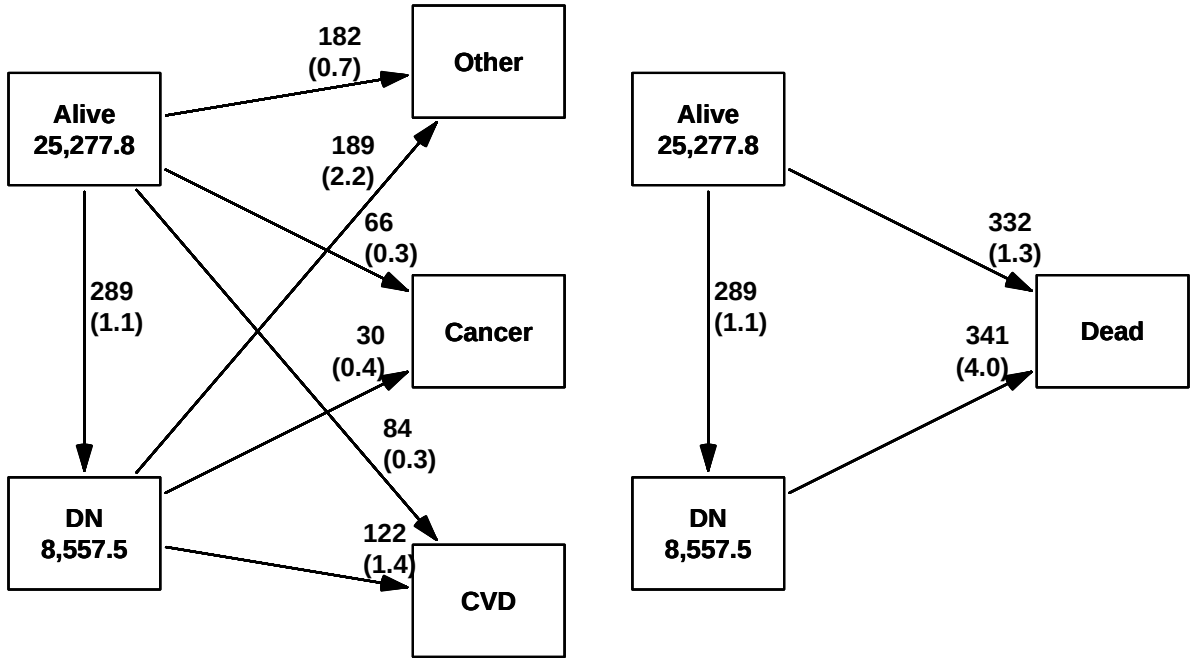

Figure 3.2: *Rates of DN and mortality rates. Number in the boxes are person-years, numbers on arrow are no. transitions (rates per 100 PY).*

3.2 Duration of complications

Now recall that the time scale `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 `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 `is.na(DNdur) & lex.Cst=="DN"`)

```
> tt <- with( C1, table(round(donef,4)) )
> table( tt )
```

```
tt
  1   2   3   6   8   9  70 629
518  30   4   1   1   2   1   1
```

```
> 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 ) )
```

```
        comp94
lex.Cst  None Prev <NA>
  Alive  3814    0    0
  DN      630  666    0
  CVD       0    0    0
  Cancer    0    0    0
  Other     0    0    0
  <NA>      0    0    0
```

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

```
                  comp94
donef              None Prev <NA>  Sum
  1984.28199863107    .    1    .    1
  1986.74880219028    .    1    .    1
  1988.08761122519    .    1    .    1
  1988.10951403149    .    1    .    1
  1989.21834360027    .    1    .    1
  2012.20260095825    1    .    .    1
  2012.33401779603    1    .    .    1
  2012.34770704997    1    .    .    1
  <NA>             3500    .    . 3500
  Sum              4444  666    . 5110
```

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):

```
> 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  3193  289  84     66   182      3814      621   25277.80      3814
  DN        0  955 122     30   189      1296      341    8557.52      1296
  Sum    3193 1244 206     96   371      5110      962   33835.32      4821
```

```
> summary( S2 )
```

```
Transitions:
     To
From     Alive    DN CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive 101455   289  84     66   182    102076      621   25277.80      3814
  DN         0 34291 122     30   189     34632      341    8557.52      1296
  Sum   101455 34580 206     96   371    136708      962   33835.32      4821
```

So we observe that the time-splitting has expanded the number of records substantially; from 5110 to 136708. To illustrate how each person contributes to the number of records, we show the records from 3 persons:

```
> options( digits=5 )
> subset( S2, lex.id %in% c(8,15,63,783), select=c(1:8,10) )
```

```
      lex.id    age     Ddur    per    DNdur  lex.dur lex.Cst lex.Xst sex
191        8 37.960  0.14811 2010.0       NA 0.039699   Alive   Alive   F
192        8 38.000  0.18781 2010.0       NA 0.250000   Alive   Alive   F
193        8 38.250  0.43781 2010.2       NA 0.250000   Alive   Alive   F
194        8 38.500  0.68781 2010.5       NA 0.250000   Alive   Alive   F
195        8 38.750  0.93781 2010.8       NA 0.246578   Alive   Alive   F
320       15 49.915  7.13911 2002.0       NA 0.250000   Alive   Alive   M
321       15 50.165  7.38911 2002.2       NA 0.250000   Alive   Alive   M
322       15 50.415  7.63911 2002.5       NA 0.250000   Alive   Alive   M
323       15 50.665  7.88911 2002.8       NA 0.250000   Alive   Alive   M
324       15 50.915  8.13911 2003.0       NA 0.250000   Alive   Alive   M
325       15 51.165  8.38911 2003.2       NA 0.250000   Alive   Alive   M
326       15 51.415  8.63911 2003.5       NA 0.250000   Alive   Alive   M
327       15 51.665  8.88911 2003.8       NA 0.250000   Alive   Alive   M
328       15 51.915  9.13911 2004.0       NA 0.250000   Alive   Alive   M
329       15 52.165  9.38911 2004.2       NA 0.250000   Alive   Alive   M
330       15 52.415  9.63911 2004.5       NA 0.250000   Alive   Alive   M
331       15 52.665  9.88911 2004.8       NA 0.250000   Alive   Alive   M
332       15 52.915 10.13911 2005.0       NA 0.250000   Alive   Alive   M
333       15 53.165 10.38911 2005.2       NA 0.113450   Alive      DN   M
334       15 53.279 10.50256 2005.4  0.00000 0.136550      DN      DN   M
335       15 53.415 10.63911 2005.5  0.13655 0.250000      DN      DN   M
336       15 53.665 10.88911 2005.8  0.38655 0.213039      DN  Cancer   M
1631      63 79.585 39.22212 2002.0       NA 0.250000   Alive   Alive   M
1632      63 79.835 39.47212 2002.2       NA 0.250000   Alive   Alive   M
1633      63 80.085 39.72212 2002.5       NA 0.250000   Alive   Alive   M
1634      63 80.335 39.97212 2002.8       NA 0.250000   Alive   Alive   M
1635      63 80.585 40.22212 2003.0       NA 0.250000   Alive   Alive   M
1636      63 80.835 40.47212 2003.2       NA 0.139459   Alive   Other   M
```

```
21929    783 65.325  1.35755 2003.2  6.83641 0.079569      DN      DN   M
21930    783 65.405  1.43712 2003.2  6.91598 0.250000      DN      DN   M
21931    783 65.655  1.68712 2003.5  7.16598 0.250000      DN      DN   M
21932    783 65.905  1.93712 2003.8  7.41598 0.250000      DN      DN   M
21933    783 66.155  2.18712 2004.0  7.66598 0.250000      DN      DN   M
21934    783 66.405  2.43712 2004.2  7.91598 0.250000      DN      DN   M
21935    783 66.655  2.68712 2004.5  8.16598 0.250000      DN      DN   M
21936    783 66.905  2.93712 2004.8  8.41598 0.250000      DN      DN   M
21937    783 67.155  3.18712 2005.0  8.66598 0.250000      DN      DN   M
21938    783 67.405  3.43712 2005.2  8.91598 0.250000      DN      DN   M
21939    783 67.655  3.68712 2005.5  9.16598 0.250000      DN      DN   M
21940    783 67.905  3.93712 2005.8  9.41598 0.250000      DN      DN   M
21941    783 68.155  4.18712 2006.0  9.66598 0.250000      DN      DN   M
21942    783 68.405  4.43712 2006.2  9.91598 0.250000      DN      DN   M
21943    783 68.655  4.68712 2006.5 10.16598 0.250000      DN      DN   M
21944    783 68.905  4.93712 2006.8 10.41598 0.250000      DN      DN   M
21945    783 69.155  5.18712 2007.0 10.66598 0.250000      DN      DN   M
21946    783 69.405  5.43712 2007.2 10.91598 0.133984      DN   Other   M
```

```
> options( digits=7 )
```

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*.

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

```
     lex.id          age             Ddur            per           DNdur
 Min.   :   1   Min.   : 9.74   Min.   : 0.00   Min.   :2002   Min.   : 0.0000
 1st Qu.:1211   1st Qu.:36.76   1st Qu.:12.76   1st Qu.:2004   1st Qu.: 0.0000
 Median :2417   Median :47.76   Median :23.19   Median :2006   Median : 0.0000
 Mean   :2415   Mean   :48.24   Mean   :24.25   Mean   :2007   Mean   : 0.5972
 3rd Qu.:3618   3rd Qu.:59.51   3rd Qu.:34.11   3rd Qu.:2009   3rd Qu.: 0.0000
 Max.   :4821   Max.   :99.60   Max.   :76.88   Max.   :2011   Max.   :16.6829
```

```
    lex.dur            lex.Cst         lex.Xst           dodth        sex           dobth
 Min.   :0.0001711   Alive :102076   Alive :101455   Min.   :2002     M:73224   Min.   :1908
 1st Qu.:0.2500000   DN    : 34632   DN    : 34580   1st Qu.:2006     F:63484   1st Qu.:1947
 Median :0.2500000   CVD   :     0   CVD   :   206   Median :2008               Median :1959
 Mean   :0.2475006   Cancer:     0   Cancer:    96   Mean   :2008               Mean   :1958
 3rd Qu.:0.2500000   Other :     0   Other :   371   3rd Qu.:2010               3rd Qu.:1970
 Max.   :0.2500000                                   Max.   :2011               Max.   :1995
                                                     NA's   :125095
      dodm           exit          entry             codth           docvd           cvd
 Min.   :1934   Min.   :2002   Min.   :2002              :125095   Min.   :1987    Min.   :1
 1st Qu.:1972   1st Qu.:2012   1st Qu.:2002   Other      :  3544   1st Qu.:1994    1st Qu.:1
 Median :1983   Median :2012   Median :2002   CVD        :  3372   Median :2001    Median :1
 Mean   :1982   Mean   :2012   Mean   :2003   Cancer     :  1699   Mean   :2001    Mean   :1
 3rd Qu.:1994   3rd Qu.:2012   3rd Qu.:2002   Infection:    812   3rd Qu.:2006    3rd Qu.:1
 Max.   :2011   Max.   :2012   Max.   :2011   Lung       :   810   Max.   :2012    Max.   :1
                                             (Other)    :  1376   NA's   :59081   NA's   :59081
      dodr             dr            donef             nef            doneu            neu
 Min.   :1983    Min.   :1       Min.   :1984    Min.   :1       Min.   :1987    Min.   :1
 1st Qu.:1994    1st Qu.:1       1st Qu.:1994    1st Qu.:1       1st Qu.:1994    1st Qu.:1
 Median :1994    Median :1       Median :1994    Median :1       Median :1994    Median :1
 Mean   :1998    Mean   :1       Mean   :1998    Mean   :1       Mean   :1998    Mean   :1
 3rd Qu.:2002    3rd Qu.:1       3rd Qu.:2003    3rd Qu.:1       3rd Qu.:2002    3rd Qu.:1
 Max.   :2012    Max.   :1       Max.   :2012    Max.   :1       Max.   :2012    Max.   :1
 NA's   :37148   NA's   :37148   NA's   :97150   NA's   :97150   NA's   :85828   NA's   :85828
    CoD           comp94
 Alive :125095   None:116526
 CVD   :  3372   Prev: 20182
 Cancer:  1699
 Other :  6542
```

Note that it is important that the changing of `NA`s 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:

```
> 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%
45.43409 58.51882 67.70705 75.45298 84.33347
```

```
> 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%
 9.821241 21.164216 31.198265 43.517833 58.434936
```

```
> 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%
2002.746 2006.322 2010.030
```

```
> 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%
1951.784 1973.646 1994.994
```

```
> 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%
 0.00000 21.41479 33.16840 43.64555
```

```
> 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)) ) )
```

```
              20%      40%      60%      80%
 0.00000 17.32976 27.48772 36.22528 46.17168
```

```
> 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 )
> cref <- 5
> ( c.kn <- with( subset(S2,lex.Cst=="DN" & lex.Xst!="DN" & comp94=="None"),
+                 c(0,quantile(DNdur+lex.dur,probs=1:3/4,na.rm=TRUE)) ) )
```

```
              25%      50%      75%
0.000000 1.980835 5.835729 8.469541
```

```
> c.pr <- seq(0,20,,n.pr)
> c.ct <- Ns(     c.pr      , knots=c.kn )
> c.rf <- Ns( rep(cref,n.pr), knots=c.kn )
```

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 )
```

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.

```
> 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) )
```

```
      sex    P  A lex.id    age     Ddur     per DNdur lex.dur lex.Cst lex.Xst dodth    dobth
33712   F 2007 12   2344 12.295 5.151007 2007.00     0    0.25   Alive   Alive    NA 1994.705
33713   F 2007 12   2344 12.545 5.401007 2007.25     0    0.25   Alive   Alive    NA 1994.705
33714   F 2007 12   2344 12.795 5.651007 2007.50     0    0.25   Alive   Alive    NA 1994.705
          dodm     exit    entry codth docvd cvd dodr dr donef nef doneu neu   CoD comp94 rate E
33712 2001.849 2012.433 2005.728          NA  NA   NA NA    NA  NA    NA  NA Alive   None    0 0
33713 2001.849 2012.433 2005.728          NA  NA   NA NA    NA  NA    NA  NA Alive   None    0 0
33714 2001.849 2012.433 2005.728          NA  NA   NA NA    NA  NA    NA  NA Alive   None    0 0
```

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

```
Classes 'Lexis' and 'data.frame':        136705 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  12 13 13 13 13 13 13 14 14 14 ...
 $ lex.id : int  1942 825 3175 825 825 825 3545 3545 3545 3545 ...
 $ age    : num  12.7 13 13.9 13.5 13.2 ...
 $ Ddur   : num  4.31 5.77 11.22 6.27 6.02 ...
 $ per    : num  2003 2002 2003 2002 2002 ...
 $ DNdur  : num  0 0 0 0 0 0 0 0 0 0 ...
 $ lex.dur: num  0.191 0.25 0.114 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  1990 1989 1989 1989 1989 ...
 $ dodm   : num  1999 1996 1992 1996 1996 ...
```

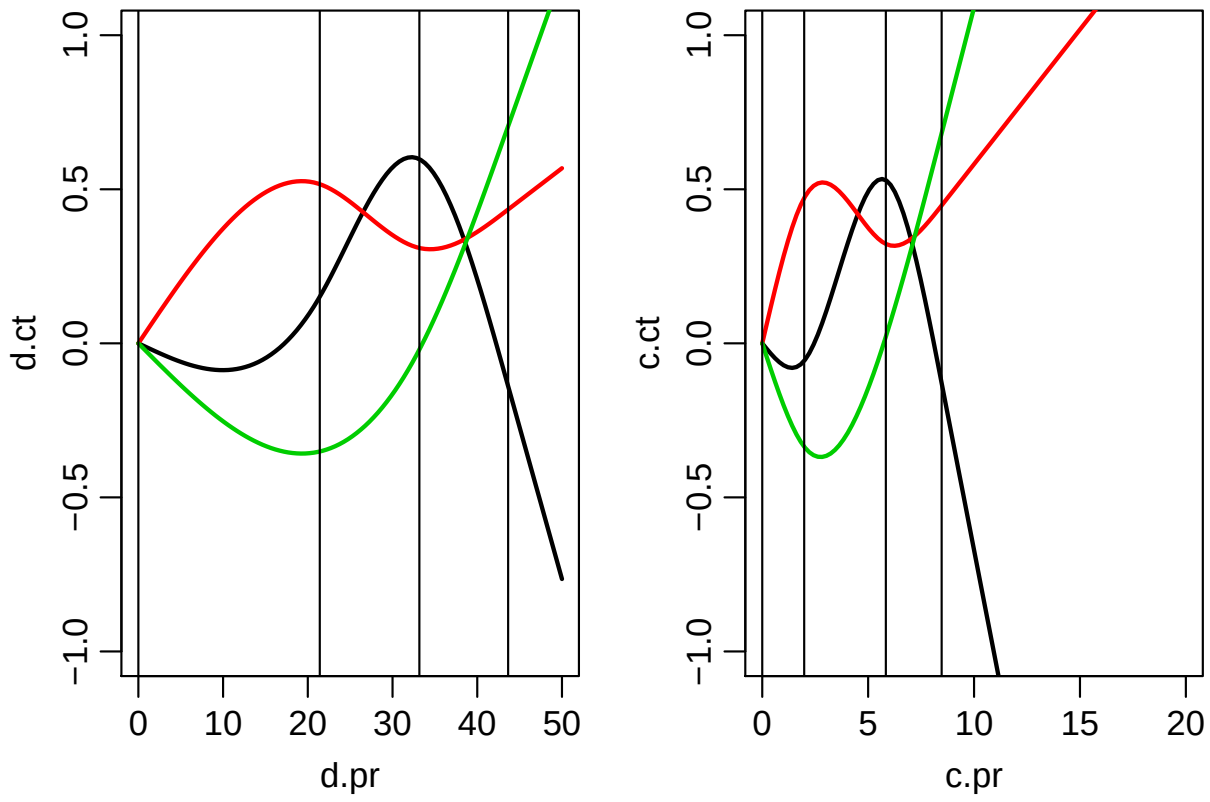

Figure 3.3: *The columns of the contrast matrices for diabetes and complications durations; note that for both duration codings, the value of all columns is 0 at 0.*

```
 $ exit   : num  2012 2012 2012 2012 2012 ...
 $ entry  : num  2003 2002 2003 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 NA NA NA NA NA ...
 $ dr     : num  NA NA NA NA NA NA NA NA NA NA ...
 $ 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 1998 NA 1998 1998 ...
 $ neu    : num  NA 1 NA 1 1 1 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  0.1256 0.0325 0.0325 0.0325 0.0325 ...
 $ E      : num  2.40e-05 8.14e-06 3.72e-06 8.14e-06 8.14e-06 ...
 - attr(*, "breaks")=List of 4
  ..$ age  : NULL
  ..$ Ddur : NULL
  ..$ per  : num  1980 1980 1980 1981 1981 ...
  ..$ DNdur: NULL
 - attr(*, "time.scales")= chr  "age" "Ddur" "per" "DNdur"
 - attr(*, "time.since")= chr  "" "" "" "DN"
```

```
> summary( S2 )
```

```
Transitions:
     To
From     Alive    DN CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive 101452   289  84     66   182    102073      621   25277.05      3814
  DN         0 34291 122     30   189     34632      341    8557.52      1296
  Sum   101452 34580 206     96   371    136705      962   33834.57      4821
```

This enables us to model the mortality rates and SMR as a function of age, calendar time, diabetes duration and complication status, and plot the rates and the RRs. But we first set up an array to hold the predicted rates and RRs, and we also make space in the array for analyses by cause of death.

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 diabetes type and sex):

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

```
      type    old   new
1  lex.Cst  Alive Alive
2  lex.Cst     DN    DN
3  lex.Cst    CVD
4  lex.Cst Cancer
5  lex.Cst  Other
6  lex.Xst  Alive Alive
7  lex.Xst     DN    DN
8  lex.Xst    CVD  Dead
9  lex.Xst Cancer  Dead
10 lex.Xst  Other  Dead
```

```
> 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 )
```

	exp(Est.)	2.5%	97.5%
(Intercept)	0.587	0.357	0.963
Ns(age, knots = a.kn)1	3.232	2.130	4.904
Ns(age, knots = a.kn)2	6.182	4.022	9.503
Ns(age, knots = a.kn)3	21.990	12.560	38.502
Ns(age, knots = a.kn)4	9.880	6.014	16.229
Ns(per, knots = p.kn)1	0.675	0.470	0.969
Ns(per, knots = p.kn)2	0.667	0.537	0.827
Ns(Ddur, knots = d.kn)1	1.010	0.450	2.265
Ns(Ddur, knots = d.kn)2	1.422	0.234	8.661
Ns(Ddur, knots = d.kn)3	0.974	0.360	2.635
I(age * Ddur/100)	1.011	0.968	1.056
I(lex.Cst == "DN")TRUE	2.378	1.938	2.918
comp94Prev	0.962	0.768	1.206

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:

`+AoD` Adding age at diagnosis to the model.

`-AoF` Removing age at follow-up from the model with both.

`-AoD` Removing age at follow-up and age at diagnosis from the model.

`+PoD` Adding date of diagnosis to the model.

`-PoF` Removing date of follow-up from the model with both.

`-PoD` Removing date of follow-up and date of diagnosis entirely.

`-comp94` Removing the indicator of prevalent complications from the base model.

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

```
> ta <- NArray( list( cod = c("AllD","SMR"),
+                      sex = levels(S2$sex),
+                     test = c("+AoD","-AoF","-AoD",
+                              "+PoD","-PoF","-PoD","-comp94"),
+                     what = 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 `A` and `P` refer to age and date of diagnosis of DM, whereas `a`, `p` and `d` refer to current age, date and diabetes duration, respectively:

```
> system.time(
+ for( rs in dimnames(ta)[["cod"]][1:2] )
+ for( sx in dimnames(ta)[["sex"]] )
+ {
+ apd <- glm( (lex.Xst=="Dead") ~ Ns( age , knots=a.kn )
+                               + Ns( per , knots=p.kn )
+                               + Ns( Ddur, knots=d.kn )
+                               + I(lex.Cst=="DN")
+                               + comp94,
+             offset = log(if(rs=="SMR") E else lex.dur/100),
+             family=poisson,
+             data = subset(S3,sex==sx) )
+ aApd <- update(  apd, . ~ . + Ns( age-Ddur, knots=A.kn ) )
+  Apd <- update( aApd, . ~ . - Ns( age     , knots=a.kn ) )
+   pd <- update(  Apd, . ~ . - Ns( age-Ddur, knots=A.kn ) )
+ apPd <- update(  apd, . ~ . + Ns( per-Ddur, knots=P.kn ) )
+  aPd <- update( apPd, . ~ . - Ns( per     , knots=p.kn ) )
+   ad <- update(  aPd, . ~ . - Ns( per-Ddur, knots=P.kn ) )
+ apdc <- update(  apd, . ~ . - comp94 )
+ ta[rs,sx,,]  <- abs( as.matrix( rbind(
+                      anova( apd,aApd,Apd,pd, test="Chisq" )[-1,c(4,3,5)],
+                      anova( apd,apPd,aPd,ad, test="Chisq" )[-1,c(4,3,5)],
+                      anova( apd,apdc,        test="Chisq" )[-1,c(4,3,5)] ) ) )
+ }
+ )
```

```
   user  system elapsed
  93.88    4.15   99.66
```

```
> round( ftable( ta[,,,3], row.vars=3 ), 3 )
```

```
        cod  AllD         SMR
        sex     M     F     M     F
test
+AoD        0.318 0.174 0.352 0.174
-AoF        0.037 0.608 0.335 0.569
-AoD        0.000 0.000 0.000 0.000
+PoD        0.958 0.935 0.975 0.924
-PoF        0.236 0.961 0.216 0.882
-PoD        0.002 0.102 0.078 0.488
-comp94     0.978 0.486 0.995 0.491
```

The conclusion from these p-values is that for men there is a strong effect of age at follow-up even in the presence of age at diagnosis but not vice-versa. 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 all sexes and

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 15, 20, 25,. . . ,50, 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 1.1.2010.

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 `NA`s, in order to get predictions that also have single `NA`s in it so that plotting of a set of separate age-curves can be done in one statement.

```
> 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,] )
```

```
    age  per Ddur lex.Cst lex.dur  E aD
1    NA   NA   NA    <NA>      NA NA NA
62 40.5 2010  0.5   Alive    1000  1 40
63 41.0 2010  1.0   Alive    1000  1 40
64 41.5 2010  1.5   Alive    1000  1 40
65 42.0 2010  2.0   Alive    1000  1 40
66 42.5 2010  2.5   Alive    1000  1 40
```

```
> new.frame <- f0[NULL,]
> for( da in seq(10,50,5) )
+ new.frame <- rbind( new.frame,
+                     make.frame(da) )
> str( new.frame )
```

```
'data.frame':        1089 obs. of  7 variables:
 $ age    : num  NA 10.5 11 11.5 12 12.5 13 13.5 14 14.5 ...
 $ 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 10 10 10 10 10 10 10 10 10 ...
```

```
> summary( new.frame )
```

```
      age             per            Ddur        lex.Cst        lex.dur           E
 Min.   :10.50   Min.   :2010   Min.   : 0.50   Alive:1080   Min.   :1000   Min.   :1
 1st Qu.:44.50   1st Qu.:2010   1st Qu.:15.38   DN   :   0   1st Qu.:1000   1st Qu.:1
 Median :60.25   Median :2010   Median :30.25   Dead :   0   Median :1000   Median :1
 Mean   :58.86   Mean   :2010   Mean   :31.64   NA's :   9   Mean   :1000   Mean   :1
 3rd Qu.:75.12   3rd Qu.:2010   3rd Qu.:46.00                3rd Qu.:1000   3rd Qu.:1
 Max.   :90.00   Max.   :2010   Max.   :80.00                Max.   :1000   Max.   :1
 NA's   :9       NA's   :9      NA's   :9                    NA's   :9      NA's   :9
       aD
 Min.   :10.00
 1st Qu.:15.00
 Median :25.00
 Mean   :27.22
 3rd Qu.:36.25
 Max.   :50.00
 NA's   :9
```

```
> # 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:

```
> AMort <- NArray( list( cod = c("All cause","SMR","CVDonly","CVD","Cancer"),
+                        mod = c("Final","Main","Int"),
+                      state = c("Alive","DN"),
+                        sex = levels(S3$sex),
+                       pred = new.frame[,"age"],
+                       what = c("Est","lo","hi") ) )
> ComplTt <- NArray( list( cod = c("All cause","SMR","CVDonly","CVD","Cancer"),
+                          sex = levels(S3$sex),
```

```
+                          pred = c("Linear DdurxAge interaction",
+                                   "Linear CxP interaction",
+                                   "Linear period effect",
+                                   "No CxP interaction",
+                                   "Compl 1994 ne Compl",
+                                   "No Compl"),
+                          what = c("Chisq","df","Pval") ) )
> ComplRR <- NArray( list( cod = c("All cause","SMR","CVDonly","CVD","Cancer"),
+                           sex = levels(S3$sex),
+                          pred = c("Compl 1994 vs. later",
+                                   "Compl later vs. None",
+                                   "Compl vs. None",
+                                   "Change / year",
+                                   "Compl vs. None (2002)",
+                                   "Compl vs. None (2010)",
+                                   "Change / year (None)",
+                                   "Change / year (Compl)"),
+                          what = c("Est","lo","hi") ) )
> p.pr <- seq(2001,2011,,30)
> p.rf <- 2010
> p.CMs <- Ns(     p.pr              , knots=p.kn ) -
+          Ns( rep(p.rf,length(p.pr)), knots=p.kn )
> CurveRR <- NArray( list( cod = c("All cause","SMR","CVDonly","CVD","Cancer"),
+                           mod = c("Final","Main","Int"),
+                         state = c("Alive","DN"),
+                           sex = levels(S3$sex),
+                          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 $(\gamma, \beta_0, \beta_1)$, then γ is the log-RR at 2010, but we additionally want the log-RR in 2002, which is:

$$\gamma + \beta_1(2002 - 2010) - \big(0 + \beta_0(2002 - 2010)\big) = \gamma + 8\beta_0 - 8\beta_1$$

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

```
> # 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
```

```
> system.time(
+ for( rs in dimnames(AMort)[["cod"]][1:2] )
+ for( sx in dimnames(AMort)[["sex"]] )
+ {
+ # 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")
+                                 + Ns( Ddur, knots=d.kn )
+                                 + 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")
+                              + Ns( per, knots=p.kn )
+                              + 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( apdc , . ~ . + comp94 )
+ apd   <- update( apdc , . ~ . - I(lex.Cst=="DN") )
+ # Tests:
+ ComplTt[rs,sx,,] <- as.matrix( abs( anova( apdcs,
+                                            apdci,
+                                            apdcl,
+                                            apdcll,
+                                            apdc,
+                                            apdcp,
+                                            apdc,
+                                            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( apdci, 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() )
+
+ zf <- predict( apdc , newdata=DN.frame, type="link", se.fit=TRUE )
+ zm <- predict( apdcs, newdata=DN.frame, type="link", se.fit=TRUE )
+ zi <- predict( apdci, newdata=DN.frame, type="link", se.fit=TRUE )
+ AMort[rs,"Final","DN",sx,,] <- exp( cbind(zf$fit,zf$se.fit) %*% ci.mat() )
+ AMort[rs,"Main" ,"DN",sx,,] <- exp( cbind(zm$fit,zm$se.fit) %*% ci.mat() )
+ AMort[rs,"Int"  ,"DN",sx,,] <- exp( cbind(zi$fit,zi$se.fit) %*% ci.mat() )
+
+ # RR by calendar time
+ CurveRR[rs,"Final","Alive",sx,,] <-
+ CurveRR[rs,"Final","DN"   ,sx,,] <- ci.exp( apdc , subset="per"        , ctr.mat=cbind(  p.pr-p.rf
+ CurveRR[rs,"Main" ,"Alive",sx,,] <- ci.exp( apdcs, subset="FALSE:Ns", ctr.mat=p.CMs )
+ CurveRR[rs,"Main" ,"DN"   ,sx,,] <- ci.exp( apdcs, subset= "TRUE:Ns", ctr.mat=p.CMs )
+
+ # Complication effects
+ ComplRR[rs,sx,"Compl 1994 vs. later",]  <- ci.exp( apdcp, subset="Prev" )
+ ComplRR[rs,sx,"Compl later vs. None",]  <- ci.exp( apdcp, subset="DN" )
+ ComplRR[rs,sx,"Compl vs. None",]        <- ci.exp( apdc , subset="DN" )
+ ComplRR[rs,sx,"Change / year",]         <- ci.exp( apdc , subset="per" )
+ ComplRR[rs,sx,c("Compl vs. None (2002)",
+                 "Compl vs. None (2010)",
+                 "Change / year (None)",
+                 "Change / year (Compl)"),] <- ci.exp( apdcll,
+                                                       subset="DN",
+                                                       ctr.mat=CMi )
+ } )
```

```
   user  system elapsed
  74.26    1.17   77.62
```

```
> ## Just to show the parameters actually extracted in the code above:
> ci.exp( apdc )
```

```
                       exp(Est.)      2.5%      97.5%
(Intercept)            1.7226143 0.7927146  3.7433398
Ns(age, knots = a.kn)1 0.5247192 0.2963638  0.9290279
Ns(age, knots = a.kn)2 0.6915768 0.4624214  1.0342915
Ns(age, knots = a.kn)3 0.3911823 0.2308246  0.6629431
Ns(age, knots = a.kn)4 0.5537255 0.3911091  0.7839550
Ns(Ddur, knots = d.kn)1 1.3074117 0.7883525  2.1682249
Ns(Ddur, knots = d.kn)2 2.0212799 0.4058837 10.0658689
Ns(Ddur, knots = d.kn)3 1.4040053 0.9957476  1.9796490
I(lex.Cst == "DN")TRUE  2.3708690 1.8601131  3.0218699
I(per - 2010)          0.9739905 0.9301730  1.0198721
```

```
> ci.exp( apdcs )
```

```
                                          exp(Est.)      2.5%     97.5%
(Intercept)                               1.9846376 0.8943990 4.4038358
Ns(age, knots = a.kn)1                    0.5237040 0.2958664 0.9269921
Ns(age, knots = a.kn)2                    0.6921639 0.4628299 1.0351336
Ns(age, knots = a.kn)3                    0.3911955 0.2307749 0.6631308
Ns(age, knots = a.kn)4                    0.5532047 0.3907509 0.7831982
Ns(Ddur, knots = d.kn)1                   1.3007085 0.7842254 2.1573424
Ns(Ddur, knots = d.kn)2                   1.9858610 0.3986338 9.8928995
Ns(Ddur, knots = d.kn)3                   1.4012807 0.9935398 1.9763551
I(lex.Cst == "DN")TRUE                    2.6352109 1.6244316 4.2749333
I(lex.Cst == "DN")FALSE:Ns(per, knots = p.kn)1 0.9774834 0.4502349 2.1221674
I(lex.Cst == "DN")TRUE:Ns(per, knots = p.kn)1  0.6699050 0.2897727 1.5487060
I(lex.Cst == "DN")FALSE:Ns(per, knots = p.kn)2 0.7401977 0.4672019 1.1727105
I(lex.Cst == "DN")TRUE:Ns(per, knots = p.kn)2  0.9883288 0.6176142 1.5815599
```

```
> ci.exp( apdcp )
```

```
                       exp(Est.)      2.5%      97.5%
(Intercept)            1.6735672 0.7657173  3.6577820
Ns(age, knots = a.kn)1 0.5244392 0.2962785  0.9283036
Ns(age, knots = a.kn)2 0.6875271 0.4593964  1.0289447
Ns(age, knots = a.kn)3 0.3886154 0.2291292  0.6591126
Ns(age, knots = a.kn)4 0.5470650 0.3857548  0.7758299
Ns(Ddur, knots = d.kn)1 1.3506554 0.8081226  2.2574173
Ns(Ddur, knots = d.kn)2 2.1153323 0.4223032 10.5957782
Ns(Ddur, knots = d.kn)3 1.4470468 1.0160734  2.0608199
I(lex.Cst == "DN")TRUE  2.5616985 1.8532197  3.5410258
I(per - 2010)          0.9721845 0.9281748  1.0182811
comp94Prev             0.8757132 0.6007539  1.2765188
```

```
> save( AMort, CurveRR, ComplRR, ComplTt, file="./data/T1AllCau.Rda" )
```

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] 0.519 0.265 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 )
```

```
                           cod All cause         SMR
                           sex         M     F     M     F
pred
Linear DdurxAge interaction        0.471 0.339 0.607 0.331
Linear CxP interaction             0.771 0.337 0.875 0.335
Linear period effect               0.233 0.962 0.212 0.883
No CxP interaction                 0.548 0.824 0.541 0.816
Compl 1994 ne Compl                0.998 0.486 0.969 0.492
No Compl                           0.000 0.000 0.000 0.000
```

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):

```
> round( ftable( ComplRR[1:2,,,], row.vars=c(3) ), 2 )
```

```
                    cod  All cause                          SMR
                    sex          M                 F          M                 F
                    what       Est   lo   hi   Est   lo   hi  Est   lo   hi  Est   lo   hi
pred
Compl 1994 vs. later          1.00 0.75 1.32 0.87 0.60 1.27 0.99 0.75 1.32 0.88 0.60 1.28
Compl later vs. None          2.18 1.68 2.83 2.57 1.86 3.55 2.19 1.69 2.84 2.56 1.85 3.54
Compl vs. None                2.18 1.78 2.67 2.37 1.86 3.02 2.18 1.78 2.67 2.37 1.86 3.02
Change / year                 0.93 0.90 0.97 0.95 0.91 1.00 0.96 0.92 0.99 0.97 0.93 1.02
Compl vs. None (2002)         1.98 1.36 2.87 2.27 1.42 3.63 1.98 1.36 2.87 2.26 1.41 3.61
Compl vs. None (2010)         2.39 1.67 3.42 2.47 1.63 3.74 2.39 1.67 3.43 2.47 1.63 3.74
Change / year (None)          0.92 0.87 0.98 0.95 0.89 1.01 0.95 0.89 1.00 0.97 0.91 1.03
Change / year (Compl)         0.94 0.90 1.00 0.96 0.89 1.03 0.97 0.92 1.02 0.98 0.92 1.05
```

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 7% (3–10) for men and a decrease of 5% (0–9) for women, whereas the corresponding decreases in SMR are M: 4% (1–8) and W: 3% (−2–7).

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.

```
> dimnames( ComplRR )
```

```
$cod
[1] "All cause" "SMR"       "CVDonly"   "CVD"       "Cancer"

$sex
[1] "M" "F"
```

```
$pred
[1] "Compl 1994 vs. later"  "Compl later vs. None"  "Compl vs. None"        "Change / year"
[5] "Compl vs. None (2002)" "Compl vs. None (2010)" "Change / year (None)"  "Change / year (Compl)"

$what
[1] "Est" "lo"  "hi"
```

```
> Est <- ComplRR[1:2,,3:8,]
> str( Est )
```

```
 num [1:2, 1:2, 1:6, 1:3] 2.179 2.182 2.373 2.371 0.934 ...
 - attr(*, "dimnames")=List of 4
  ..$ cod : chr [1:2] "All cause" "SMR"
  ..$ sex : chr [1:2] "M" "F"
  ..$ pred: chr [1:6] "Compl vs. None" "Change / year" "Compl vs. None (2002)" "Compl vs. None (2010
  ..$ what: chr [1:3] "Est" "lo" "hi"
```

```
> Est[,,c(2,5,6),] <- round((Est[,,c(2,5,6),]-1)*100,1)
> round( ftable( Est, col.vars=c(2,4), row.vars=c(3,1) ), 1 )
```

		sex	M			F		
		what	Est	lo	hi	Est	lo	hi
pred	cod							
Compl vs. None	All cause		2.2	1.8	2.7	2.4	1.9	3.0
	SMR		2.2	1.8	2.7	2.4	1.9	3.0
Change / year	All cause		-6.6	-10.1	-3.0	-4.8	-9.0	-0.3
	SMR		-4.3	-7.9	-0.6	-2.6	-7.0	2.0
Compl vs. None (2002)	All cause		2.0	1.4	2.9	2.3	1.4	3.6
	SMR		2.0	1.4	2.9	2.3	1.4	3.6
Compl vs. None (2010)	All cause		2.4	1.7	3.4	2.5	1.6	3.7
	SMR		2.4	1.7	3.4	2.5	1.6	3.7
Change / year (None)	All cause		-7.7	-12.7	-2.5	-5.2	-10.9	0.9
	SMR		-5.5	-10.6	-0.1	-3.1	-8.9	3.1
Change / year (Compl)	All cause		-5.5	-10.3	-0.5	-4.2	-10.6	2.6
	SMR		-3.2	-8.1	2.0	-2.0	-8.5	4.9

From this we see that there is a significant decrease of 5% per year for male T1D and a 3% non-significant decrease in mortalty among women.

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

```
> # The age at diagnosis
> pref <- 2010
> pr.A <- as.numeric( dimnames(AMort)[["pred"]] )
> agr <- cumsum( is.na(pr.A) )
> wh <- agr %in% c(2,5,8) ## Age at dx 15, 30, 45
> 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(15,90),
+       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):

```
> 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(15,90),
+       xlab="Age", ylab="", las=1 )
> abline( v=seq(0,100,5), h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
> abline( h=1 )
> 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 ), 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:

```
> rlim <- c(1/2,5)
> p.pr <- as.numeric( dimnames( CurveRR )[["pred"]] )
```

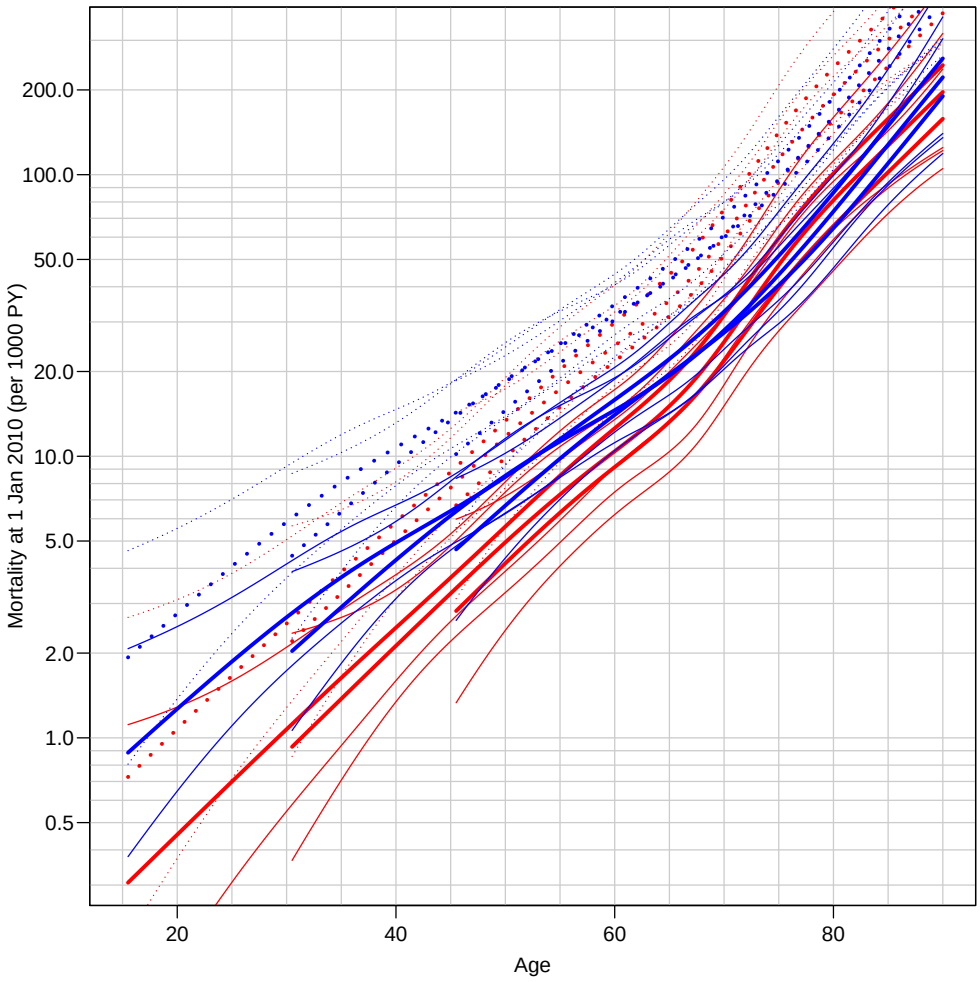

Figure 3.4: *Age-specific all-cause mortality rates without (full lines) and with (broken lines) DN for T1 patients aged 15. 30, 45 at diagnosis (assumed to be in 2010). Red: F; blue: M; thin lines: 95% c.i.*

```
> 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) )
> abline( h=1 )
> for( sx in 2:1 ) # sx <- 1
+ matlines( p.pr, cbind( CurveRR["All cause","Final","Alive",sx,,],
+                         CurveRR["All cause","Main" ,"Alive",sx,,],
+                         CurveRR["All cause","Final","DN"   ,sx,,],
+                         CurveRR["All cause","Main" ,"DN"   ,sx,,] ),
+           lwd=c(3,1,1), lty=rep(c(1,3),each=6), col=c("blue","red")[sx] )
> mtext( "Mortality RR", line=2.5, side=2, outer=FALSE )
> box()
```

To explore the shape of the (non-significant) interactions we plot all the mortality curves

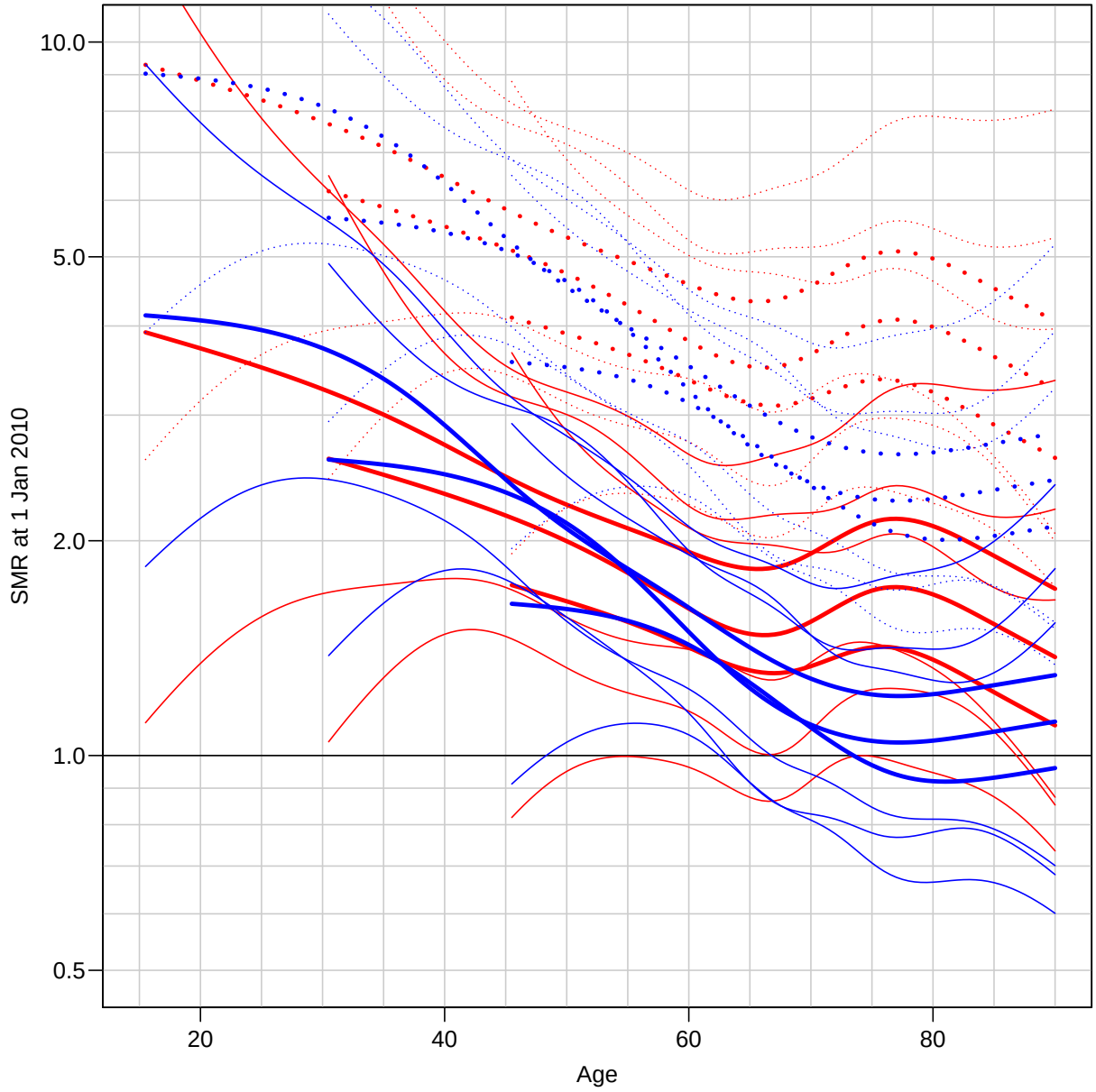

Figure 3.5: *Age-specific all-cause relative mortality (SMR). T1 patients aged 15, 30, 45 at diagnosis, T2 patients age 40, 55, 70 at diagnosis. Red: F; blue: M; thin lines: 95% c.i.*

for the two models together:

```
> par( mfrow=c(2,3), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, oma=c(0,2,2,0) )
> 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(10,90), 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

```
> str( ComplRR )
```

```
 num [1:5, 1:2, 1:8, 1:3] 1 0.995 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 [1:8] "Compl 1994 vs. later" "Compl later vs. None" "Compl vs. None" "Change / year"
  ..$ what: chr [1:3] "Est" "lo" "hi"
```

```
> arttab <- ComplRR[1:2,,,]
> dimnames( arttab )
```

```
$cod
[1] "All cause" "SMR"

$sex
[1] "M" "F"

$pred
[1] "Compl 1994 vs. later"  "Compl later vs. None"  "Compl vs. None"        "Change / year"
[5] "Compl vs. None (2002)" "Compl vs. None (2010)" "Change / year (None)"  "Change / year (Compl)"

$what
[1] "Est" "lo"  "hi"
```

```
> arttab[,,c(4,7,8),] <- round( ( arttab[,,c(4,7,8),] - 1 )*100, 1 )
> round( ftable( arttab, row.vars=c(3,1)), 2)
```

		sex	M			F	
		what	Est	lo	hi	Est	lo	hi
pred	cod						
Compl 1994 vs. later	All cause		1.00	0.75	1.32	0.87	0.60	1.27
	SMR		0.99	0.75	1.32	0.88	0.60	1.28
Compl later vs. None	All cause		2.18	1.68	2.83	2.57	1.86	3.55
	SMR		2.19	1.69	2.84	2.56	1.85	3.54
Compl vs. None	All cause		2.18	1.78	2.67	2.37	1.86	3.02
	SMR		2.18	1.78	2.67	2.37	1.86	3.02
Change / year	All cause		-6.60	-10.10	-3.00	-4.80	-9.00	-0.30
	SMR		-4.30	-7.90	-0.60	-2.60	-7.00	2.00
Compl vs. None (2002)	All cause		1.98	1.36	2.87	2.27	1.42	3.63
	SMR		1.98	1.36	2.87	2.26	1.41	3.61
Compl vs. None (2010)	All cause		2.39	1.67	3.42	2.47	1.63	3.74

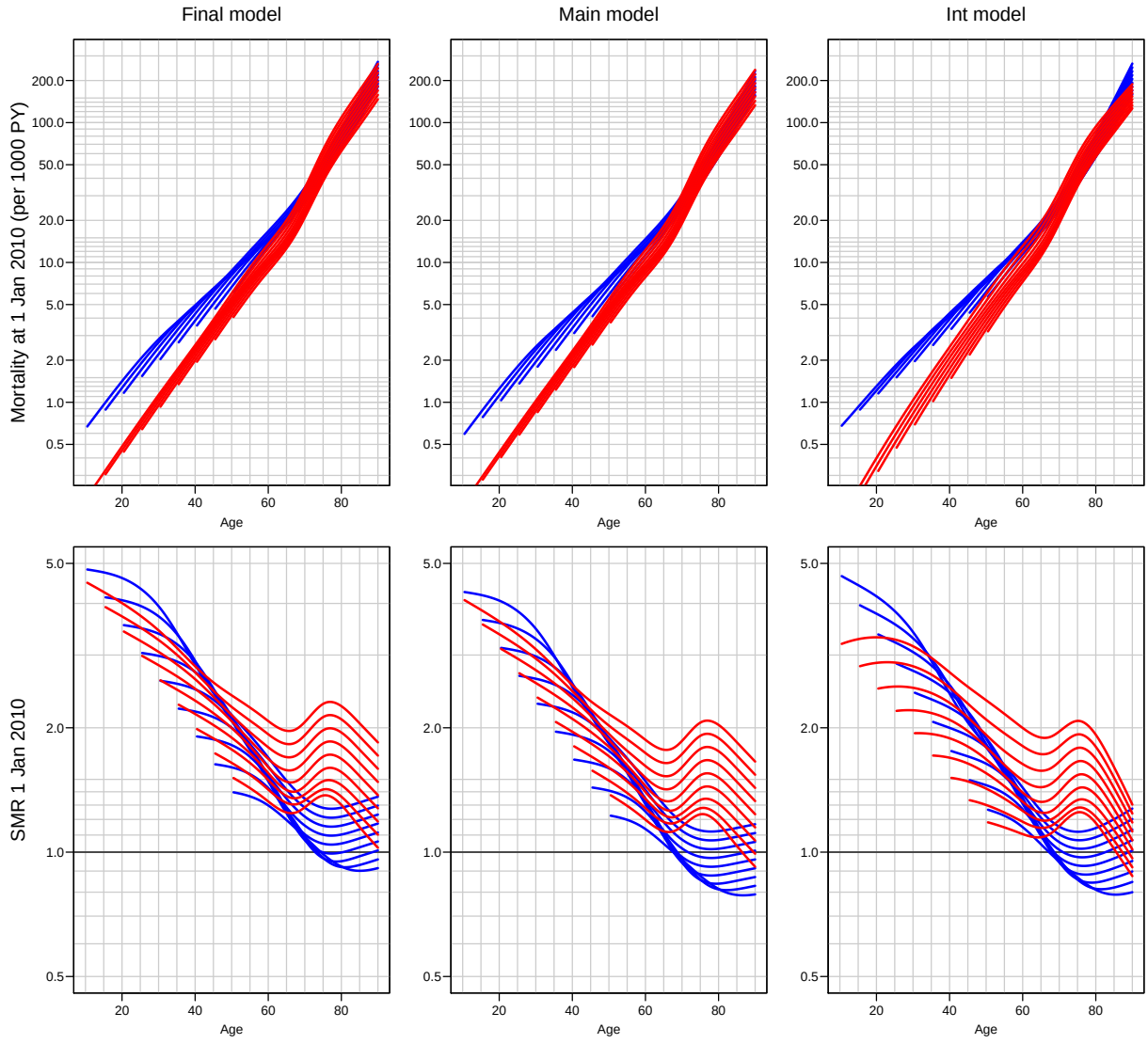

Figure 3.6: *Age-specific all-cause mortality and SMR for T1D patients without complications, diagnosed 1.1.2010 in ages 10, 20, 30 and 40. The final model have a linear effect of calendar time, the main and the interaction model a 2-parameter-spline effect of calendar time; in all models the calendar 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*

```
                      SMR              2.39   1.67   3.43   2.47   1.63   3.74
Change / year (None)  All cause       -7.70 -12.70  -2.50  -5.20 -10.90   0.90
                      SMR             -5.50 -10.60  -0.10  -3.10  -8.90   3.10
Change / year (Compl) All cause       -5.50 -10.30  -0.50  -4.20 -10.60   2.60
                      SMR             -3.20  -8.10   2.00  -2.00  -8.50   4.90
```

It is the last 6 lines of output from each diabetes type that enters as the table in the paper.

We then draw the figure(s) for the paper:

```
> f1 <-
+ function(cls=1:3,dr=TRUE,my=0.55,ry=1.4)
+ {
+ par( mfrow=if( dr ) c(3,2) else c(2,2), mar=c(3,0,1,0), oma=c(0,4,1,1), mgp=c(3,1,0)/1.6,
+      las=1, lend=1 )
+ scol <- c("blue","red")
+ pr.A <- as.numeric( dimnames(AMort)[["pred"]] )
+ agr <- cumsum( is.na(pr.A) )
+ wh <- agr %in% c(2,5,8)   ## Age at dx 15, 30, 45
+ fs <- which(diff(wh)>0)+2 ## First point of predictions
+
+ p.pr <- as.numeric( dimnames(CurveRR)[["pred"]] )
+
+ ######################################################################
+ # Mortality
+ mlim <- c(1,900)/3
+ for( sx in 1:2 )
+    {
+ plot( NA, type="n", log="y", ylim=mlim, xlim=c(15,90),
+       xlab="Age", ylab="", las=1, yaxt="n" )
+ abline( v=seq(0,100,5), h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
+ matlines( pr.A[wh], AMort["All cause","Final","Alive",sx,wh,cls],
+           lwd=c(3,1,1), lty=1, col=scol[sx] )
+ matlines( pr.A[wh], AMort["All cause","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("Mortality at 1 Jan", pref, "(per 1000 PY)"),
+          line=2.5, side=2, outer=FALSE, las=0, cex=0.7 )
+   }
+ if( !is.null(my) )
+   {
+   segments( floor(pr.A[fs]),
+             pmin(my,apply(AMort["All cause","Final",c("Alive","DN"),sx,fs,1],2,min)),
+             floor(pr.A[fs]),
+             pmax(my,apply(AMort["All cause","Final",c("Alive","DN"),sx,fs,1],2,max)),
+             col=scol[sx] )
+   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(15,90),
+       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( paste("Mortality/SMR ratio", pref ),
+          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()
> pdf( "T1Fig1.pdf", height=9, width=7 )
> f1( cls=1 )
> dev.off()


pdf
  2


> f1( cls=1 )
```

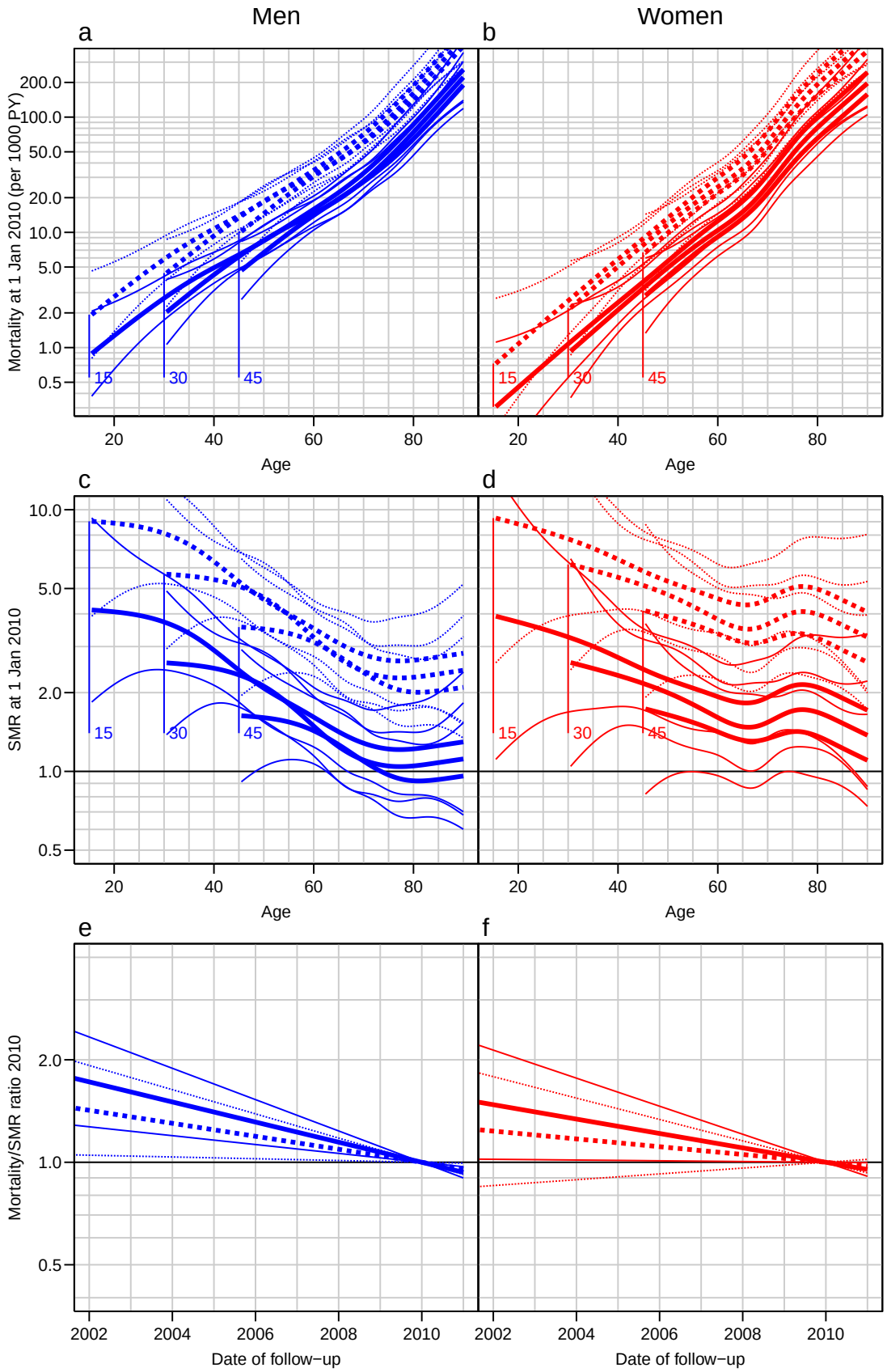

Figure 3.7: *Mortality rates (a,b), SMR (c,d) and changes in these (e,f). Patients diagnosed with DM in ages 15, 30 and 45. For mortality rates and SMR, patients without nephropathy are with full lines, patients with nephropathy are with dotted lines. For the changes (e,f), the full lines are changes in mortality, dotted lines changes in SMR. Thin lines indicate 95% confidence intervals.*

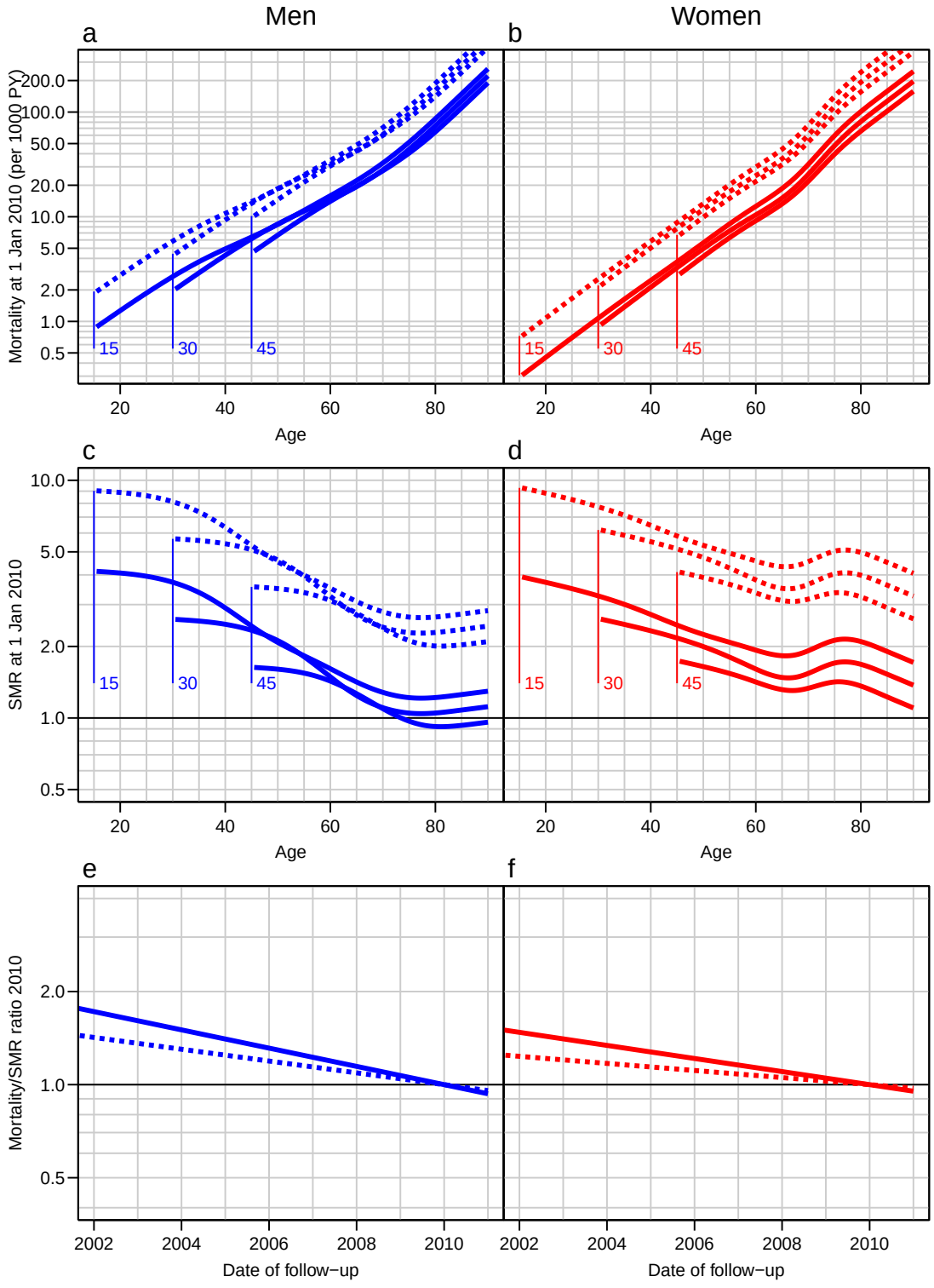

Figure 3.8: *Mortality rates (a,b), SMR (c,d) and changes in these (e,f). Patients diagnosed with DM in ages 15, 30 and 45. For mortality rates and SMR, patients without nephropathy are with full lines, patients with nephropathy are with dotted lines. For the changes (e,f), the full lines are changes in mortality, dotted lines changes in SMR.*

Bibliography