Mortality among Type 2 patients at Steno Diabetes Center

SDC
http://BendixCarstensen.com/SDC/EPJmort
July 2013
Version 2.2

Compiled Monday 7th October, 2013, 21:54
from: C:/Bendix/Steno/MaEJ/EPJ-dod/MortT2.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 2 patients only at Steno Diabetes Center in the period where the computerized patients records (EPJ) are available, that is the period 1.9.2001–15.9.2010.

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

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

1.2 Data preparation

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

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

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



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


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

---------------------------------------------------------------
C:\Bendix\Steno\MaEJ\EPJ-dod\sas\getit.sas
---------------------------------------------------------------
NOTE: Libref HER was successfully assigned as follows:
      Engine:        V9
      Physical Name: C:\Bendix\Steno\MaEJ\EPJ-dod\sas
NOTE: Libref DATA was successfully assigned as follows:
      Engine:        V9
      Physical Name: 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_1.1.57     foreign_0.8-53

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

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

```
> 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=="T2", select=-2 )
> str( epj )


'data.frame':        5876 obs. of  15 variables:
 $ dodth: num  15651 NA 17860 NA NA ...
 $ sex  : num  2 2 1 1 1 2 2 1 1 2 ...
 $ dobth: num  -15340 -13149 -12053 -11322 -10957 ...
 $ dodm : num  6210 11323 10958 11323 9862 ...
 $ exit : num  15651 19151 17860 19151 19151 ...
 $ entry: num  15063 15063 15063 15371 15063 ...
 $ codth: Factor w/ 10 levels "","accidents",..: 8 1 10 1 1 1 1 5 1 1 ...
 $ docvd: num  12329 14879 14199 17405 13482 ...
 $ cvd  : num  1 1 1 1 1 NA 1 1 1 NA ...
 $ dodr : num  12329 NA 14199 NA 17157 ...
 $ dr   : num  1 NA 1 NA 1 1 NA 1 NA NA ...
 $ donef: num  12329 12329 15498 18852 12329 ...
 $ nef  : num  1 1 1 1 1 NA NA 1 1 NA ...
 $ doneu: num  12329 12329 14199 NA 12329 ...
 $ neu  : num  1 1 1 NA 1 NA 1 1 1 1 ...
```

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
1 2002.85   F  1918 1977.54 2002.85 2001.24 kidney 1993.75   1 1993.75  1 1993.75   1 1993.75   1
2      NA   F  1924 1991.33 2012.43 2001.24         2000.74   1      NA NA 1993.75   1 1993.75   1
3 2008.90   M  1927 1990.18 2008.90 2001.24  other 1998.87   1 1998.87  1 2002.43   1 1998.87   1
4      NA   M  1929 1991.92 2012.43 2002.08         2007.65   1      NA NA 2011.61   1       NA  NA
5      NA   M  1930 1987.76 2012.43 2001.24         1996.91   1 2006.97  1 1993.75   1 1993.75   1
6      NA   F  1930 1988.78 2012.43 2004.96              NA  NA 2004.96  1       NA  NA       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          18    0    0    0    0    0    0    0    0    0   18
    2002          71    1    0    0    0    0    0    0    0    0   72
    2003          72    7    8    0    0    0    0    0    0    0   87
    2004          78    4    7    0    0    0    0    0    0    0   89
    2005          70   12    7    3    3    0    0    0    0    0   95
    2006          59   11    7    7    8    1    0    0    0    0   93
    2007          42    9    6    3    8    5    3    0    0    0   76
    2008          54   10    7    4    5    3    6    3    0    0   92
    2009          65    4    8    5    3    5    4    7    1    0  102
    2010          59    7    4    6    2    7    8    2    0    1   96
    NA           939  157  177  240  189  167  182  165  203  202 2621
    Sum         1527  222  231  268  218  188  203  177  204  203 3441
F   2001          14    0    0    0    0    0    0    0    0    0   14
    2002          46    4    0    0    0    0    0    0    0    0   50
    2003          38    2    2    0    0    0    0    0    0    0   42
    2004          55    1    8    1    0    0    0    0    0    0   65
    2005          45    6    3    4    1    0    0    0    0    0   59
    2006          49    6    9    9    4    1    0    0    0    0   78
    2007          44    1    3    7    1    1    0    0    0    0   57
    2008          46    2    2    6    4    0    5    1    0    0   66
    2009          32    7    1    8    6    2    2    4    3    0   65
    2010          35    2    3    3    6    4    1    0    3    0   57
    NA           717  113  131  130  142  136  136  127  129  121 1882
```

```
    Sum          1121  144  162  168  164  144  144  132  135  121 2435
Sum 2001           32    0    0    0    0    0    0    0    0    0   32
    2002          117    5    0    0    0    0    0    0    0    0  122
    2003          110    9   10    0    0    0    0    0    0    0  129
    2004          133    5   15    1    0    0    0    0    0    0  154
    2005          115   18   10    7    4    0    0    0    0    0  154
    2006          108   17   16   16   12    2    0    0    0    0  171
    2007           86   10    9   10    9    6    3    0    0    0  133
    2008          100   12    9   10    9    3   11    4    0    0  158
    2009           97   11    9   13    9    7    6   11    4    0  167
    2010           94    9    7    9    8   11    9    2    3    1  153
    NA           1656  270  308  370  331  303  318  292  332  323 4503
    Sum          2648  366  393  436  382  332  347  309  339  324 5876
```

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 2010-12-31 So we just check if all persons enter and exits correctly:

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

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

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

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

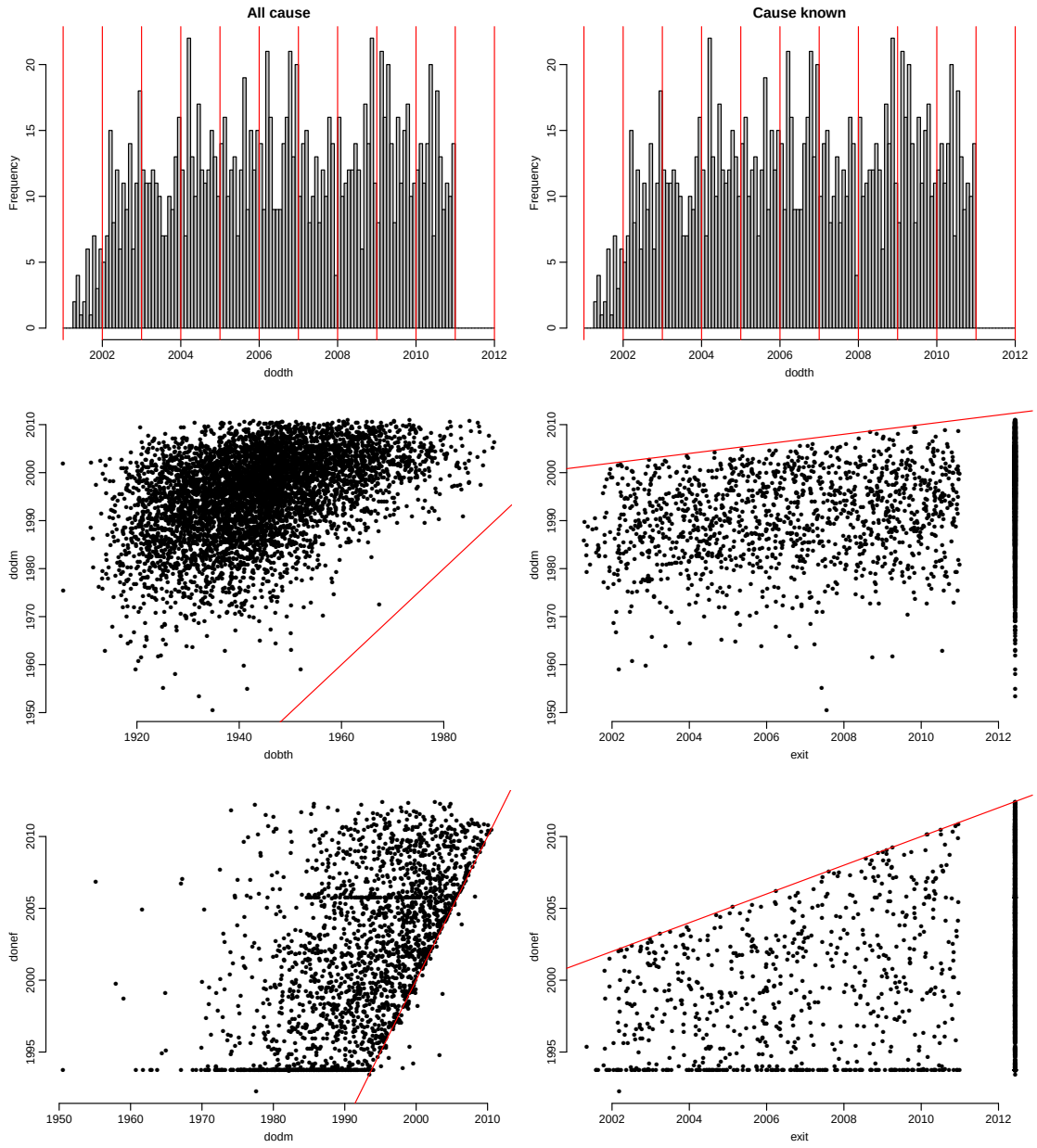

Figure 1.1: *Histogram of dates of death for all known deaths and for deaths where a cause is known. The bottom 4 panels have plots of the date of diagnosis resp. complications versus date of birth and date of exit, with a red line indicating the identity (so all points should preferably be on the same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

```
          Dead
DN         FALSE TRUE <NA>  Sum
  1993.755   114  157    0  271
  1994.001     2    5    0    7
  2001.999     7    1    0    8
  2005.747    48    4    0   52
  2005.749     7    1    0    8
  <NA>      3073  632    0 3705
  Sum       4503 1373    0 5876
```

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
     4503        17        11       288       493        67        91        34        84       288
```

```
> # 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
             4503    0      0     0
  CVD           0  493      0     0
  Cancer        0    0    288     0
  Acute DM      0    0      0    11
  Kidney        0    0      0    34
  Lung          0    0      0    84
  GI            0    0      0    67
  Infection     0    0      0    91
  Accidents     0    0      0    17
  Other         0    0      0   288
```

```
> 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   11      3     4   18     0    9      0     5   14
2002          0   23     11    38   72     0   25      7    18   50
2003          0   41     17    29   87     0   16     12    14   42
2004          0   31     18    40   89     0   22     14    29   65
2005          0   39     25    31   95     0   27      8    24   59
2006          0   32     17    44   93     0   33     15    30   78
2007          0   20     21    35   76     0   14     11    32   57
2008          0   32     19    41   92     0   24     12    30   66
2009          0   32     24    46  102     0   21     16    28   65
2010          0   26     22    48   96     0   15     16    26   57
NA         2621    0      0     0 2621  1882    0      0     0 1882
Sum        2621  287    177   356 3441  1882  206    111   236 2435
```

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  2621  820       3441       820   20144.39       3441

$F

Transitions:
     To
From    Alive Dead  Records:  Events: Risk time:  Persons:
  Alive  1882  553       2435       553   14597.89       2435
```

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 142246 1373    143619     1373   34742.28      5876
```

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	12.4	0.0	0.0	9.1	0.0
2001.25	3.0	330.5	0.9	4.0	239.9	1.7
2001.5	5.0	347.1	1.4	4.0	254.6	1.6
2001.75	10.0	367.6	2.7	6.0	267.7	2.2
2002	17.0	383.6	4.4	10.0	280.9	3.6
2002.25	15.0	394.7	3.8	11.0	288.1	3.8
2002.5	22.0	400.4	5.5	12.0	293.3	4.1
2002.75	18.0	410.3	4.4	17.0	298.1	5.7
2003	17.0	420.2	4.0	17.0	303.0	5.6
2003.25	25.0	430.4	5.8	8.0	312.2	2.6
2003.5	19.0	437.5	4.3	5.0	318.8	1.6
2003.75	26.0	447.5	5.8	12.0	327.1	3.7
2004	27.0	457.3	5.9	14.0	334.7	4.2
2004.25	28.0	468.5	6.0	12.0	343.1	3.5
2004.5	14.0	477.5	2.9	21.0	347.6	6.0
2004.75	20.0	489.7	4.1	18.0	353.4	5.1
2005	23.0	501.0	4.6	17.0	358.2	4.7
2005.25	23.0	508.7	4.5	9.0	366.1	2.5
2005.5	24.0	515.2	4.7	16.0	371.3	4.3
2005.75	25.0	522.7	4.8	17.0	379.5	4.5
2006	25.0	531.4	4.7	19.0	385.5	4.9
2006.25	21.0	537.1	3.9	13.0	391.0	3.3
2006.5	19.0	542.9	3.5	20.0	394.6	5.1
2006.75	28.0	548.0	5.1	26.0	398.3	6.5
2007	26.0	553.6	4.7	13.0	402.1	3.2
2007.25	13.0	562.2	2.3	18.0	407.8	4.4
2007.5	19.0	569.3	3.3	11.0	410.7	2.7
2007.75	19.0	577.8	3.3	15.0	416.1	3.6
2008	21.0	585.5	3.6	16.0	423.5	3.8
2008.25	21.0	590.4	3.6	17.0	425.5	4.0
2008.5	18.0	593.9	3.0	17.0	429.1	4.0
2008.75	31.0	599.7	5.2	16.0	435.5	3.7
2009	27.0	606.8	4.4	18.0	438.6	4.1
2009.25	29.0	613.8	4.7	13.0	442.3	2.9
2009.5	22.0	618.0	3.6	20.0	449.1	4.5
2009.75	24.0	625.0	3.8	14.0	453.6	3.1
2010	23.0	633.1	3.6	14.0	456.4	3.1
2010.25	29.0	639.3	4.5	12.0	460.2	2.6
2010.5	23.0	644.4	3.6	17.0	464.0	3.7
2010.75	21.0	649.4	3.2	14.0	467.6	3.0

```
> matplot( as.numeric(dimnames(DY)[[1]]), log="y", las=1,
+          xlab="Date", ylab="Raw mortality (% / year)",
+          DY[,,"rate"], type="l", lty=1, lwd=3, col=c("blue","red") )
> abline( v=seq(1998,2015,1/4), col=gray(0.9) )
> abline( v=seq(1998,2015,1)  , col=gray(0.8) )
> box()
```

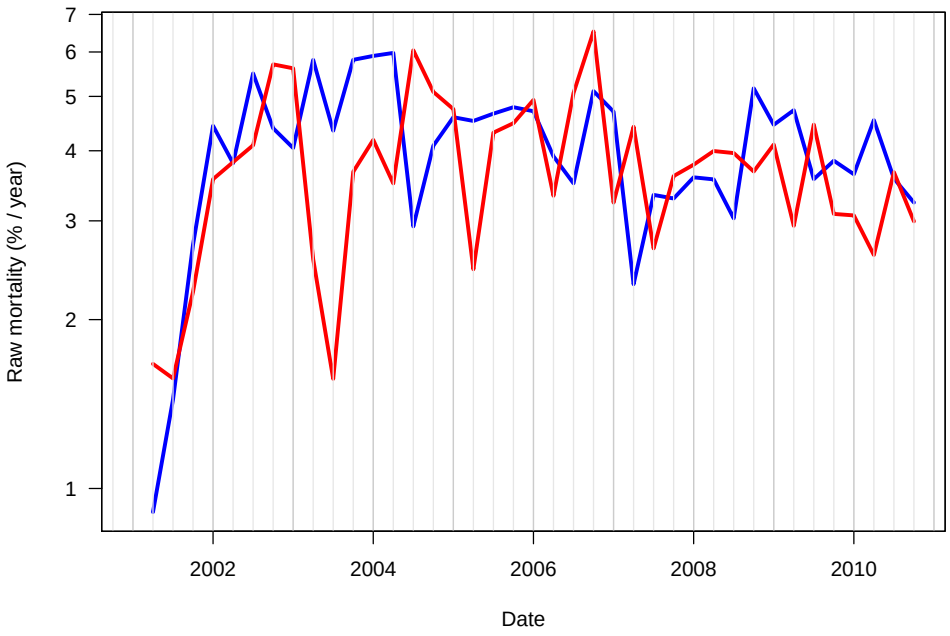

Figure 1.2: *Raw mortality rates for T2 patients by 3-month periods; there is something missing prior to 2002.*

A quick look at the tables or figure 1.2 shows that data seem incomplete prior to 2002, so to make sure that data are valid, we define entry to be at the start of 2002, and redefine the Lexis object:

```
> 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" )

$M

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

$F

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

> summary.data.frame( L1 )

      age              Ddur              per            lex.dur           lex.Cst      lex.Xst
 Min.   :14.290   Min.   : 0.0000   Min.   :2002.0   Min.   :0.0068446   Alive:5844   Alive:4503
 1st Qu.:52.389   1st Qu.: 2.3210   1st Qu.:2002.0   1st Qu.:2.9226557   Dead :   0   Dead :1341
 Median :61.073   Median : 6.9664   Median :2002.8   Median :6.0479124
 Mean   :60.392   Mean   : 8.6015   Mean   :2004.4   Mean   :5.6320081
 3rd Qu.:69.454   3rd Qu.:12.8486   3rd Qu.:2006.5   3rd Qu.:8.9993155
 Max.   :96.523   Max.   :51.4427   Max.   :2011.0   Max.   :8.9993155

     lex.id          dodth        sex          dobth             dodm             exit
 Min.   :   1.0   Min.   :2002.0   M:3423   Min.   :1905.6   Min.   :1950.6   Min.   :2002.0
 1st Qu.:1461.8   1st Qu.:2004.5   F:2421   1st Qu.:1934.4   1st Qu.:1990.5   1st Qu.:2012.4
 Median :2922.5   Median :2006.7            Median :1943.6   Median :1997.3   Median :2012.4
 Mean   :2922.5   Mean   :2006.7            Mean   :1944.0   Mean   :1995.8   Mean   :2011.1
 3rd Qu.:4383.2   3rd Qu.:2008.9            3rd Qu.:1952.3   3rd Qu.:2002.3   3rd Qu.:2012.4
 Max.   :5844.0   Max.   :2011.0            Max.   :1989.9   Max.   :2011.0   Max.   :2012.4
                  NA's   :4503
     entry             codth          docvd            cvd            dodr             dr
 Min.   :2002.0          :4503   Min.   :1988.6   Min.   :1      Min.   :1991.5   Min.   :1
 1st Qu.:2002.0   CVD    : 473   1st Qu.:1996.8   1st Qu.:1      1st Qu.:1996.6   1st Qu.:1
 Median :2002.8   Other  : 286   Median :2001.3   Median :1      Median :2001.9   Median :1
 Mean   :2004.4   Cancer : 285   Mean   :2001.3   Mean   :1      Mean   :2001.5   Mean   :1
 3rd Qu.:2006.5   Infection:  89   3rd Qu.:2005.1   3rd Qu.:1      3rd Qu.:2005.7   3rd Qu.:1
 Max.   :2011.0   Lung   :  82   Max.   :2012.4   Max.   :1      Max.   :2012.4   Max.   :1
                  (Other) : 126   NA's   :1045     NA's   :1045   NA's   :2879     NA's   :2879
     donef            nef           doneu            neu            CoD
 Min.   :1992.2   Min.   :1      Min.   :1991.5   Min.   :1      Alive :4503
 1st Qu.:1997.2   1st Qu.:1      1st Qu.:1996.4   1st Qu.:1      CVD   : 473
 Median :2001.9   Median :1      Median :2001.1   Median :1      Cancer: 285
 Mean   :2001.8   Mean   :1      Mean   :2000.8   Mean   :1      Other : 583
 3rd Qu.:2005.9   3rd Qu.:1      3rd Qu.:2004.6   3rd Qu.:1
 Max.   :2012.4   Max.   :1      Max.   :2012.4   Max.   :1
 NA's   :3686     NA's   :3686   NA's   :2950     NA's   :2950
```

```
> save( L1, file="./data/T2L1.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(0,190)
> with( subset( L1, sex=="M" ),
+       hist( age, breaks=0:101, col="blue", border="blue",
+             main="", ylim=yl, xlab="", ylab="" ) )
> with( subset(L1, sex=="F"),
+       hist( age, breaks=0:100, col="red", border="red",
+             main="", ylim=yl, xlab="", ylab="" ) )
> zz <-
+ with( subset( L1, sex=="M" ),
+       hist( Ddur, breaks=0:101, col="blue", border="blue",
+             main="", ylim=yl, xlab="", ylab="" ) )
> text( 1.5, yl[2], zz$counts[1], font=2, adj=c(-0.1,0), col="blue" )
> zz <-
+ with( subset(L1, sex=="F" ),
+       hist( Ddur, breaks=0:100, col="red", border="red",
+             main="", ylim=yl, xlab="", ylab="" ) )
> text( 1.5, yl[2], zz$counts[1], font=2, adj=c(-0.1,0), col="red" )
> mtext( c("Age at entry","DM duration at entry"), side=3, line=0,
+        at=c(1,3)/4, outer=TRUE )
```

1.4 `Lexis` object for analysis of cause-specific mortality

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

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

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

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

```
       entryOK
deathOK TRUE
   TRUE 1373
   <NA> 4503
```

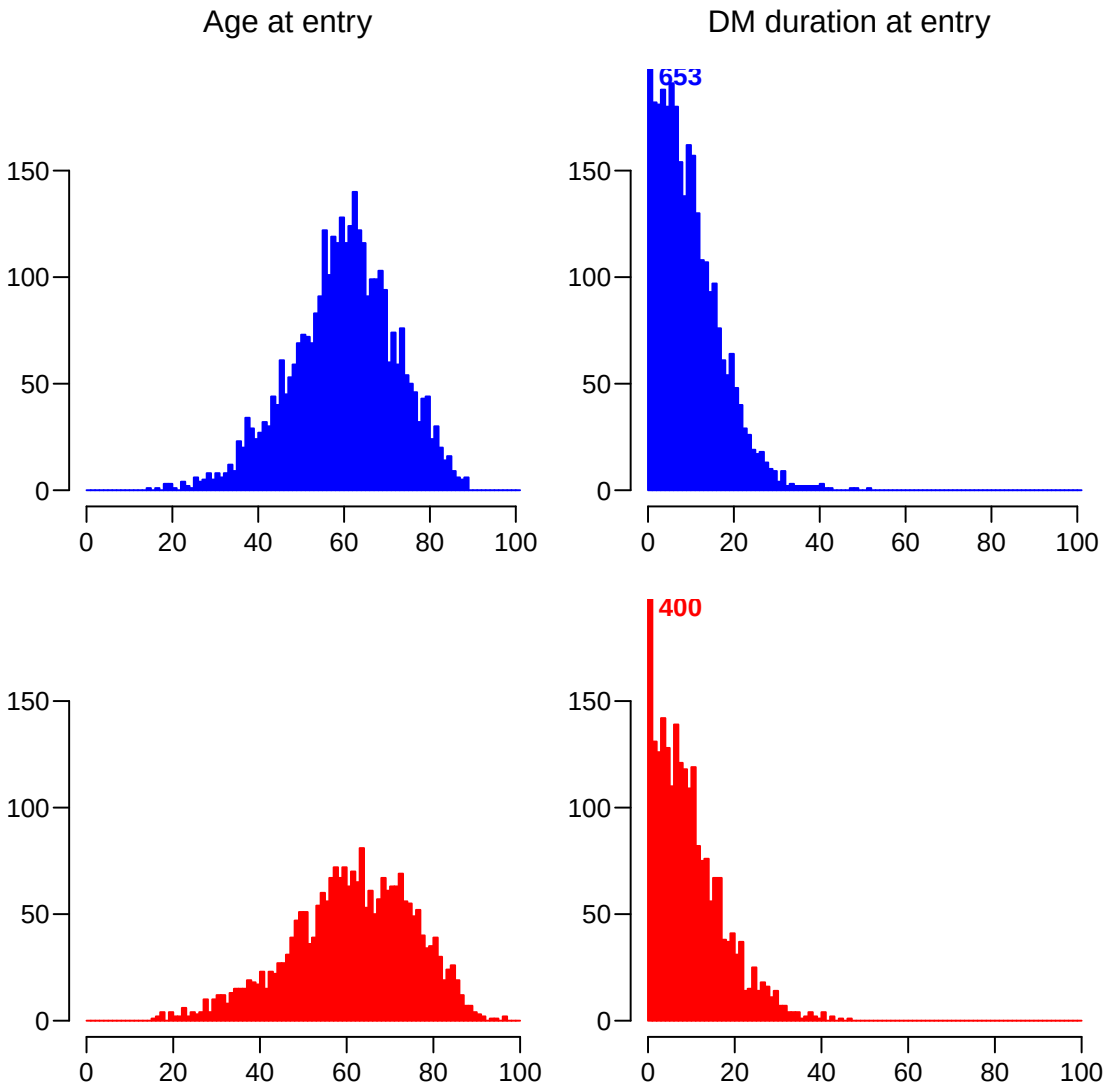

Figure 1.3: *Entry age and duration for T2 patients at Steno.*

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

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

```
> nrow( C1 )
```

```
[1] 5844
```

```
> summary( C1 )
```

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

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

```
          sex     M                              F
          CoD Alive  CVD Cancer Other  Sum Alive  CVD Cancer Other  Sum
codth
               2621    0      0     0 2621  1882    0      0     0 1882
CVD               0  276      0     0  276     0  197      0     0  197
Cancer            0    0    174     0  174     0    0    111     0  111
Acute DM          0    0      0     5    5     0    0      0     6    6
Kidney            0    0      0    21   21     0    0      0    13   13
Lung              0    0      0    50   50     0    0      0    32   32
GI                0    0      0    45   45     0    0      0    21   21
Infection         0    0      0    52   52     0    0      0    37   37
Accidents         0    0      0     7    7     0    0      0     8    8
Other             0    0      0   172  172     0    0      0   114  114
Sum            2621  276    174   352 3423  1882  197    111   231 2421
```

```
> save( C1, file="./data/T2CoD.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:

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

```
 logi [1:2, 1:12, 1:2] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 3
  ..$ sex: chr [1:2] "M" "F"
  ..$    : chr [1:12] "N" "Age" "Age-IQR" "DM dur" ...
  ..$    : chr [1:2] "N" "%/sd"
```

```
> tab1[,"N","N"] <- with( L1, table(sex) )
> tab1["M","N","%/sd"] <- tab1["M","N","N"]/(tab1["M","N","N"]+tab1["F","N","N"])*100
> tab1["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[,"FU time","N"]       <- with( L1, tapply(lex.dur,sex,median) )
> tab1[,"FU time","%/sd"]    <- with( L1, tapply(lex.dur,sex,sd) )
> tab1[,"FUtime-IQR","N"]    <- with( L1, tapply(lex.dur,sex,quantile,probs=1/4) )
> tab1[,"FUtime-IQR","%/sd"] <- with( L1, tapply(lex.dur,sex,quantile,probs=3/4) )
> tab1[,"Nephropathy","N"] <- with( L1, table((entry+0.5)>donef,sex )["TRUE",] )
> tab1[,"Neuropathy" ,"N"] <- with( L1, table((entry+0.5)>doneu,sex )["TRUE",] )
> tab1[,"Retinopathy","N"] <- with( L1, table((entry+0.5)>dodr ,sex )["TRUE",] )
> tab1[,"CVD"        ,"N"] <- with( L1, table((entry+0.5)>docvd,sex )["TRUE",] )
> tab1[,"Deaths"     ,"N"] <- with( L1, table(lex.Xst=="Dead"  ,sex )["TRUE",] )
> tab1[,8:12,"%/sd"] <- tab1[,8:12,"N"] / tab1["N",rep(1,5)] * 100
> round( ftable(tab1,col.vars=c(1,3)), 1 )
```

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

N               3423.0   58.6 2421.0   41.4
Age               61.0   12.0   62.5   14.1
Age-IQR           52.9   68.6   52.8   72.1
DM dur             7.3    7.4    7.8    7.9
DMdur-IQR          2.6   13.1    3.1   13.6
FU time            6.0    3.0    6.1    3.0
FUtime-IQR         2.8    9.0    3.1    9.0
Nephropathy     1173.0   34.3  566.0   23.4
Neuropathy      1613.0   47.1  963.0   39.8
Retinopathy     1426.0   41.7  983.0   40.6
CVD             2423.0   70.8 1630.0   67.3
Deaths           802.0   23.4  539.0   22.3
```

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

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

N               3423   59 2421   41
Age               61   12   62   14
Age-IQR           53   69   53   72
DM dur             7    7    8    8
DMdur-IQR          3   13    3   14
FU time            6    3    6    3
FUtime-IQR         3    9    3    9
Nephropathy     1173   34  566   23
Neuropathy      1613   47  963   40
Retinopathy     1426   42  983   41
CVD             2423   71 1630   67
Deaths           802   23  539   22
```

Chapter 2

Mortality by cause of death

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

```
Transitions:
     To
From     Alive CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive 132598 473    285   583    133939     1341   32913.46      5844
```

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

```
[1] 4.75 7.25
```

```
> 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(0,60), ylim=c(0,100), xaxs="i", yaxs="i",
+         xlab="Diabetes duration", ylab="Age" )
> points( S1, pch=16, cex=0.5,
+         col=c("transparent",rainbow(3))[S1$lex.Xst] )
> rect( 48,0,60,17, col="white", border="lightgray" )
> text( rep(58,5), 1:4*3.5, c(levels(S1$lex.Xst)[2:4],"C.o.D."),
+       col=c(rainbow(3),"gray"), adj=1, cex=0.9, font=2 )
> box()
```

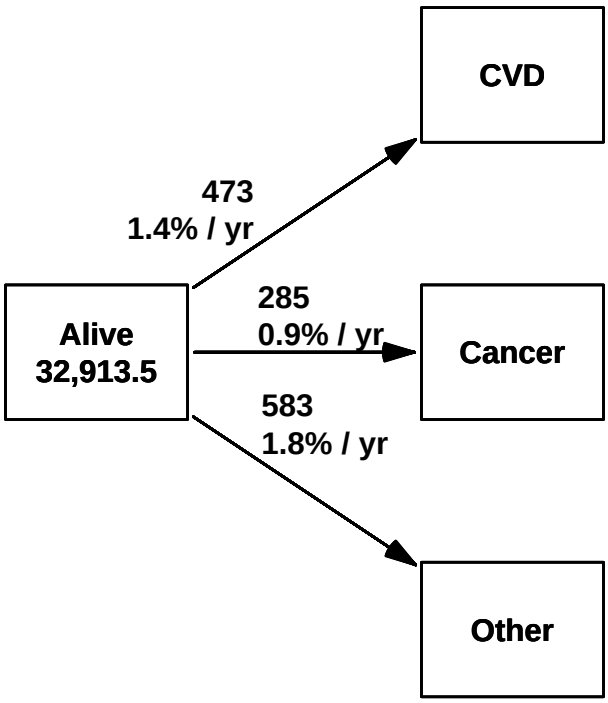

Figure 2.1: *Person-years, deaths and mortality rates by cause of death in the SDC T2 patient population.*

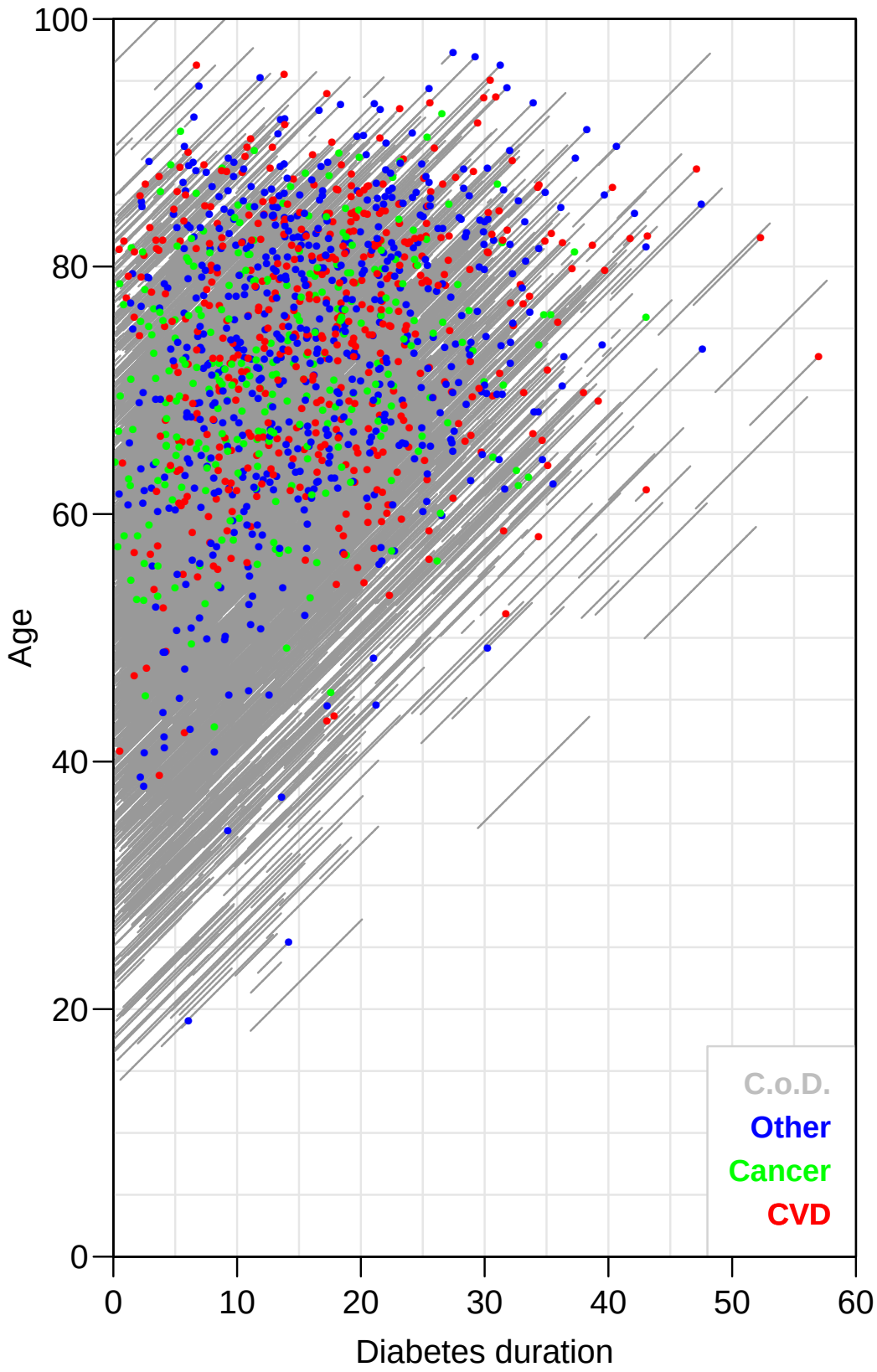

Figure 2.2: *Distribution of follow-up and deaths (by cause) for T2. Although not visible directly, no person in this plot has a life-line (total follow-up) of more than 7 years, since the earliest entry is 1.1.2002, and the latest exit is 31.12.2008.*

2.1 Statistical analysis

2.1.1 Setup

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

```
> n.pr <- 100
> ( a.kn <- with( subset(S1,lex.Xst!="Alive"),
+                 quantile(age+lex.dur,probs=c(1,3,5,7,9)/10) ) )

      10%       30%       50%       70%       90%
59.641342 67.638604 74.168378 79.989049 85.806982

> a.pr <- seq(40,95,,n.pr)
> a.ct <- Ns( a.pr, knots=a.kn )
> pref <- 2010
> ( p.kn <- with( subset(S1,lex.Xst!="Alive"),
+                 quantile(per+lex.dur,probs=c(1,5,9)/10) ) )

      10%       50%       90%
2003.0883 2006.7187 2010.1300

> p.pr <- seq(2002,2011,,n.pr)
> p.ct <- Ns(     p.pr      , knots=p.kn )
> p.rf <- Ns( rep(pref,n.pr), knots=p.kn )
> dref <- 10
>  ( d.kn <- with( subset(S1,lex.Xst!="Alive"),
+                  c(0,quantile(Ddur+lex.dur,probs=1:2/3,na.rm=TRUE)) ) )

           33.333333% 66.666667%
  0.000000  11.363240  19.399553

> d.pr <- seq(0,40,,n.pr)
> d.ct <- Ns(     d.pr      , knots=d.kn )
> d.rf <- Ns( rep(dref,n.pr), knots=d.kn )
```

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

```
> 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
  31.45    2.81   35.61
```

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

```
           what P(lin) RR/year    lo    hi P(null)
cod    sex
CVD    M         0.541   0.902 0.861 0.944   0.000
       F         0.356   0.910 0.861 0.961   0.001
Cancer M         0.621   0.981 0.926 1.040   0.526
       F         0.743   1.006 0.935 1.083   0.863
Other  M         0.325   0.959 0.920 0.999   0.045
       F         0.058   0.990 0.941 1.041   0.692
```

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

```
           what RR/year    lo    hi
cod    sex
CVD    M           -9.8 -13.9  -5.6
       F           -9.0 -13.9  -3.9
Cancer M           -1.9  -7.4   4.0
       F            0.6  -6.5   8.3
Other  M           -4.1  -8.0  -0.1
       F           -1.0  -5.9   4.1
```

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

We can now plot the age-effects:

```
> 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.

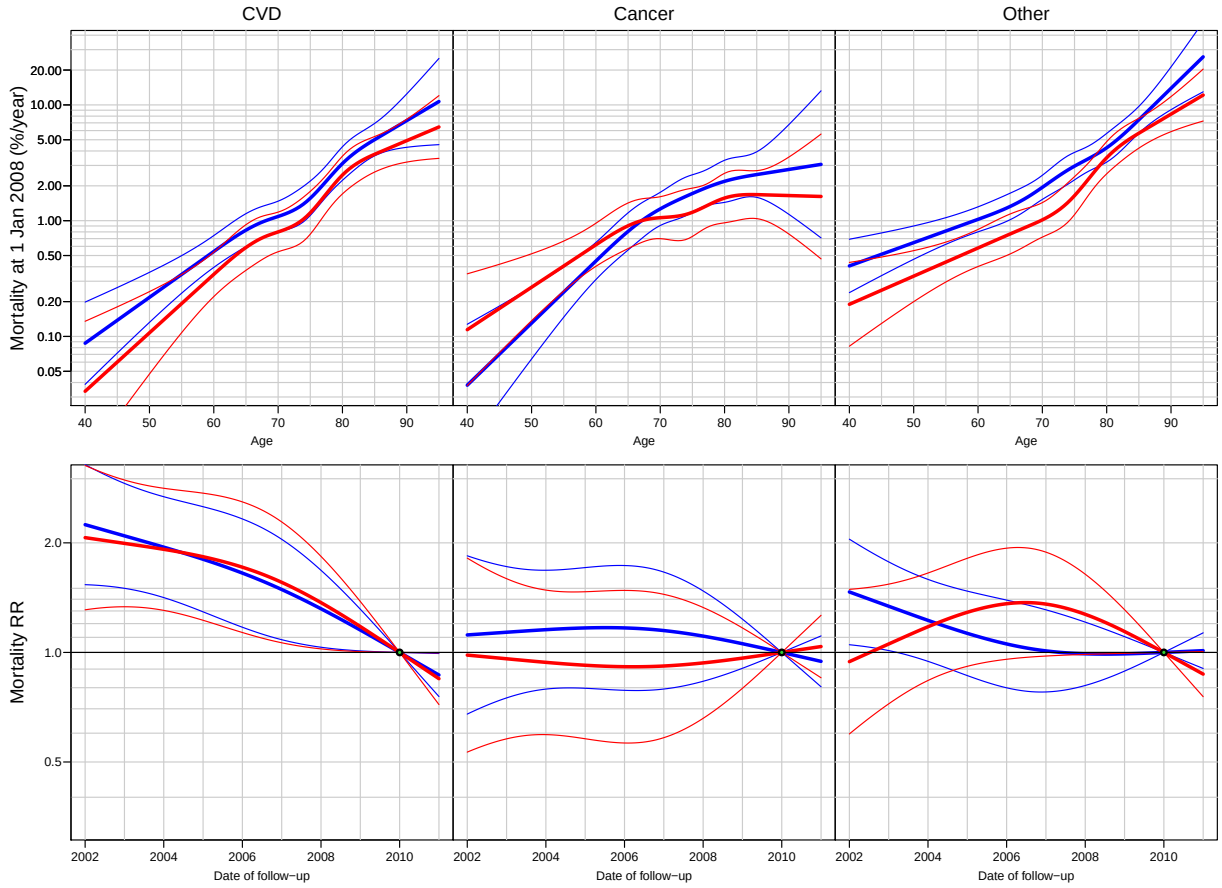

Figure 2.3: *Age-specific mortality and the change of this over calendar time for T2 diabetes patients. Red: F; blue: M, thin lines: 95% c.i.*

2.3 Including diabetes duration

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

```
> 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","DoDM"),
+                         what=c("P(lin)","RR/year","lo","hi") )) )
> str( linx )
```

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

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

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

```
   user  system elapsed
  63.88    1.42   67.33
```

We can now inspect the tests for linearity of period and duration effects as well as the estimates of the slope of the linear effects under the null:

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

```
            sex       M                          F
            what P(lin) RR/year    lo    hi P(lin) RR/year    lo    hi
cod    eff
CVD    PRR        0.466   0.897 0.857 0.940  0.303   0.904 0.856 0.955
       DRR        0.498   1.033 1.019 1.047  0.989   1.027 1.012 1.042
       DoDM          NA   0.964 0.945 0.983     NA   0.974 0.951 0.997
Cancer PRR        0.642   0.983 0.927 1.042  0.695   1.014 0.942 1.092
       DRR        0.364   0.996 0.977 1.015  0.272   0.981 0.958 1.004
       DoDM          NA   1.009 0.989 1.030     NA   1.026 1.001 1.052
Other  PRR        0.411   0.952 0.914 0.992  0.048   0.985 0.936 1.036
       DRR        0.052   1.037 1.025 1.049  0.247   1.011 0.997 1.025
       DoDM          NA   0.952 0.935 0.970     NA   0.980 0.960 1.002
```

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

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

```
           eff      PRR                 DRR                DoDM
           what RR/year    lo    hi RR/year    lo    hi RR/year    lo    hi
cod    sex
CVD    M          -10.3 -14.3  -6.0     3.3   1.9   4.7    -3.6  -5.5  -1.7
       F           -9.6 -14.4  -4.5     2.7   1.2   4.2    -2.6  -4.9  -0.3
Cancer M           -1.7  -7.3   4.2    -0.4  -2.3   1.5     0.9  -1.1   3.0
       F            1.4  -5.8   9.2    -1.9  -4.2   0.4     2.6   0.1   5.2
Other  M           -4.8  -8.6  -0.8     3.7   2.5   4.9    -4.8  -6.5  -3.0
       F           -1.5  -6.4   3.6     1.1  -0.3   2.5    -2.0  -4.0   0.2
```

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 +3%/year by duration, −10%/year by calendar time. There is no effect for Cancer, and for Other causes there seems to be a small effect by duration, in the vicinity of 4%/year, but only for men.

```
> 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 )
+     }
+  }
```

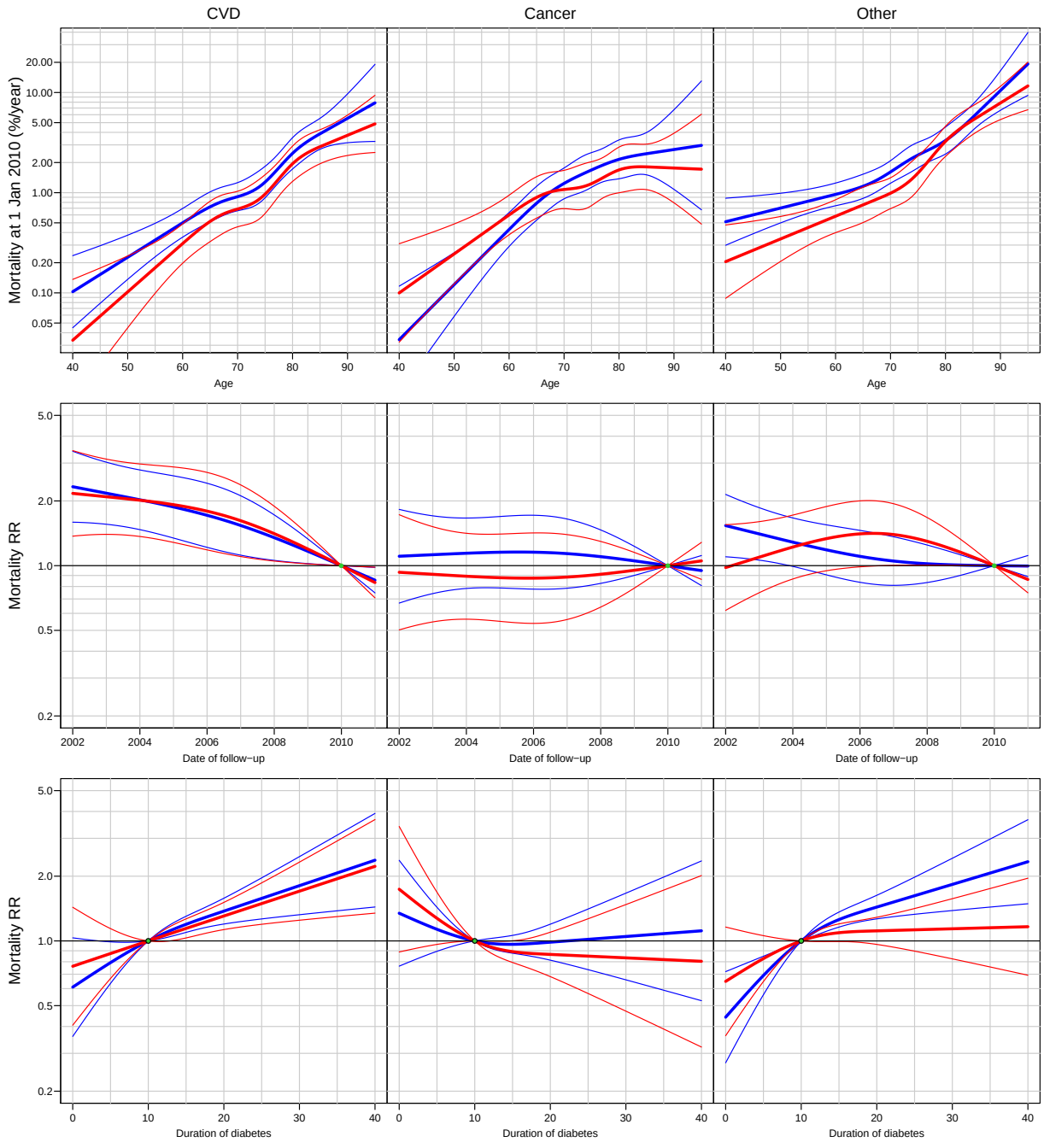

Figure 2.4: *Age-specific mortality and the change of this over calendar time for T2 patients. Red: F; blue: M, thin lines: 95% c.i.*

Chapter 3

All-cause mortality by complication status

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

```
> load( file="./data/T2L1.Rda" )
> load( file="./data/T2CoD.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          1658  112     81   131   244   25     15    34  1902  137     96   165
    TRUE              0    0      0     0   719  139     78   187   719  139     78   187
F   FALSE          1415   94     78   117   134   16      2    29  1549  110     80   146
    TRUE              0    0      0     0   333   87     31    85   333   87     31    85
Sum FALSE          3073  206    159   248   378   41     17    63  3451  247    176   311
    TRUE              0    0      0     0  1052  226    109   272  1052  226    109   272
```

3.1 Setup

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

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

```
$M

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

```
  DN        0  930 164     93   221      1408      478      75.28      1408
  Sum    1691 1215 276    174   352      3708     1087     190.87      3423

$F

Transitions:
     To
From    Alive  DN CVD Cancer Other  Records:  Events: Risk time:  Persons:
  Alive  1431 165  94     78   117      1885      454     101.40      1885
  DN        0 451 103     33   114       701      250      36.87       701
  Sum    1431 616 197    111   231      2586      704     138.27      2421


> 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
> al <- c(0,100)
> dl <- c(0,60)
```

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

```
> pdf( "./graph/T2mort-Lexis-dur.pdf", height=diff(al)/ypi+1, width=diff(dl)/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=dl, ylim=al, 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$lex.Xst] )
> rect( 50,0,60,20, col="white", border="lightgray" )
> text( rep(58,5), 1:4*3.5, c(levels(C2$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
```

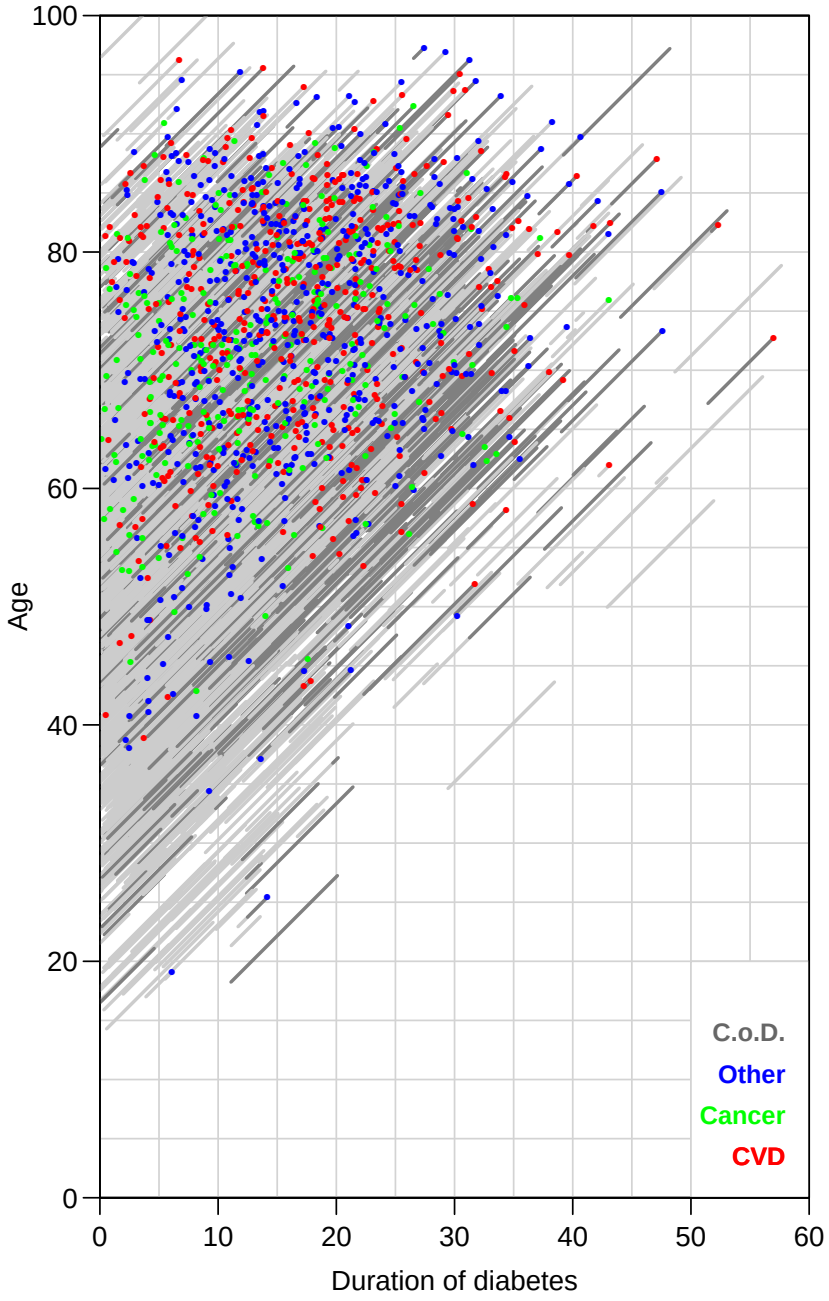

Figure 3.1: *Follow-up for T2 patients at Steno. Follow-up after onset of DN is shown in dark gray color, deaths shown as dots.*

```
> 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.

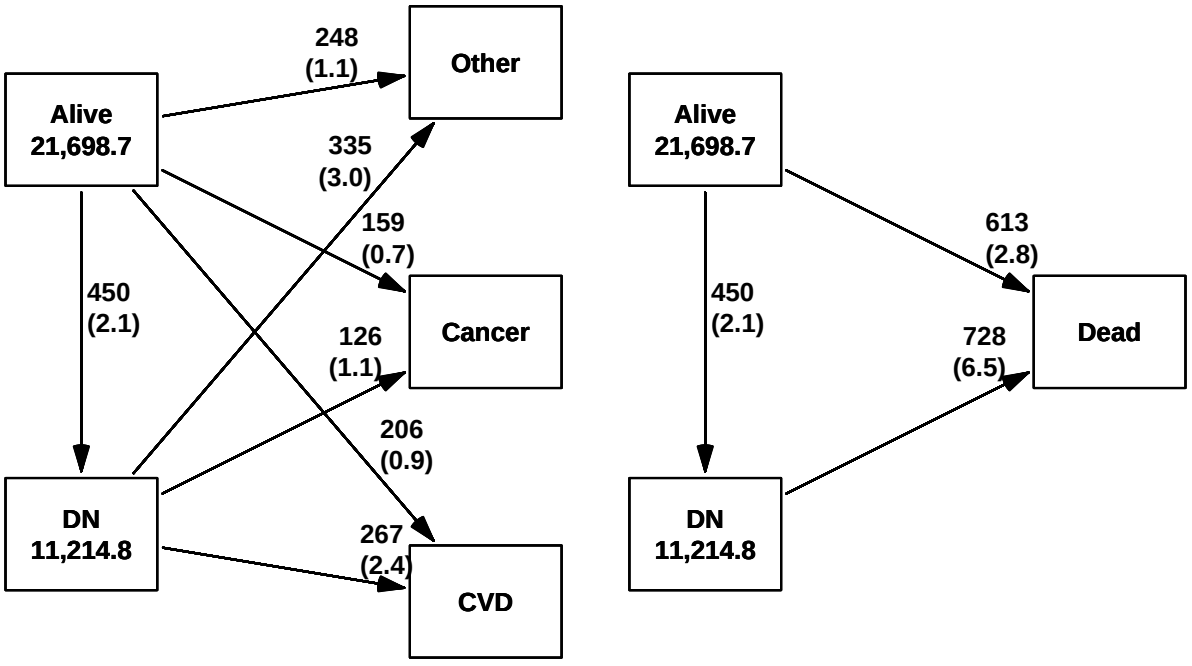

Figure 3.2: *Rates of DN and mortality rates. Number in the boxes are person-years, numbers on arrow are no. transitions (rates per 100 PY).*

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    4    5    7    8   52  265
1167  254   34    9    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  4185    0    0
  DN     1811  298    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
  1992.23545516769    .    1    .    1
  1993.41546885695    .    1    .    1
  1993.66461327858    .    1    .    1
  1993.75496235455    .  265    .  265
  1993.75770020534    .    1    .    1
  2012.34496919918    1    .    .    1
  2012.37782340862    1    .    .    1
  2012.41067761807    1    .    .    1
  <NA>             3686    .    . 3686
  Sum              5996  298    . 6294
```

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  3122  450 206    159   248      4185     1063   21698.70      4185
  DN        0 1381 267    126   335      2109      728   11214.76      2109
  Sum    3122 1831 473    285   583      6294     1791   32913.46      5844
```

```
> summary( S2 )
```

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

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

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

```
     lex.id    age     Ddur    per    DNdur   lex.dur lex.Cst lex.Xst sex
1300     47 80.060  1.37301 2005.1       NA 0.1028405   Alive   Alive   M
1301     47 80.163  1.47585 2005.2       NA 0.2500000   Alive   Alive   M
1302     47 80.413  1.72585 2005.5       NA 0.2500000   Alive   Alive   M
1303     47 80.663  1.97585 2005.8       NA 0.2500000   Alive   Alive   M
1304     47 80.913  2.22585 2006.0       NA 0.2500000   Alive   Alive   M
1305     47 81.163  2.47585 2006.2       NA 0.2500000   Alive   Alive   M
1306     47 81.413  2.72585 2006.5       NA 0.2500000   Alive   Alive   M
1307     47 81.663  2.97585 2006.8       NA 0.2500000   Alive   Alive   M
1308     47 81.913  3.22585 2007.0       NA 0.2361396   Alive     CVD   M
2591    108 84.752 14.39061 2002.0       NA 0.2500000   Alive   Alive   F
2592    108 85.002 14.64061 2002.2       NA 0.2500000   Alive   Alive   F
2593    108 85.252 14.89061 2002.5       NA 0.2500000   Alive   Alive   F
2594    108 85.502 15.14061 2002.8       NA 0.2500000   Alive   Alive   F
2595    108 85.752 15.39061 2003.0       NA 0.1813826   Alive      DN   F
2596    108 85.933 15.57199 2003.2 0.000000 0.0686174      DN      DN   F
2597    108 86.002 15.64061 2003.2 0.068617 0.2500000      DN      DN   F
2598    108 86.252 15.89061 2003.5 0.318617 0.2500000      DN      DN   F
2599    108 86.502 16.14061 2003.8 0.568617 0.0912047      DN     CVD   F
2914    125 54.672  0.73141 2008.9       NA 0.0800821   Alive   Alive   F
2915    125 54.752  0.81149 2009.0       NA 0.2500000   Alive   Alive   F
2916    125 55.002  1.06149 2009.2       NA 0.2500000   Alive   Alive   F
2917    125 55.252  1.31149 2009.5       NA 0.2500000   Alive   Alive   F
2918    125 55.502  1.56149 2009.8       NA 0.2500000   Alive   Alive   F
2919    125 55.752  1.81149 2010.0       NA 0.2500000   Alive   Alive   F
2920    125 56.002  2.06149 2010.2       NA 0.2500000   Alive   Alive   F
2921    125 56.252  2.31149 2010.5       NA 0.2500000   Alive   Alive   F
2922    125 56.502  2.56149 2010.8       NA 0.2493155   Alive   Alive   F
3121    133 67.670 29.41918 2002.0 1.917864 0.2500000      DN      DN   F
```

```
3122    133 67.920 29.66918 2002.2 2.167864 0.2500000      DN      DN   F
3123    133 68.170 29.91918 2002.5 2.417864 0.2500000      DN      DN   F
3124    133 68.420 30.16918 2002.8 2.667864 0.2500000      DN      DN   F
3125    133 68.670 30.41918 2003.0 2.917864 0.2500000      DN      DN   F
3126    133 68.920 30.66918 2003.2 3.167864 0.2500000      DN      DN   F
3127    133 69.170 30.91918 2003.5 3.417864 0.2500000      DN      DN   F
3128    133 69.420 31.16918 2003.8 3.667864 0.2500000      DN      DN   F
3129    133 69.670 31.41918 2004.0 3.917864 0.0027379      DN   Other   F
```

```
> 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.   : 14.29   Min.   : 0.000   Min.   :2002   Min.   : 0.0000
 1st Qu.:1445   1st Qu.: 55.27   1st Qu.: 6.068   1st Qu.:2005   1st Qu.: 0.0000
 Median :2912   Median : 63.52   Median :10.890   Median :2007   Median : 0.0000
 Mean   :2916   Mean   : 62.87   Mean   :12.233   Mean   :2007   Mean   : 1.5128
 3rd Qu.:4379   3rd Qu.: 71.85   3rd Qu.:16.925   3rd Qu.:2009   3rd Qu.: 0.9227
 Max.   :5844   Max.   :100.93   Max.   :57.397   Max.   :2011   Max.   :16.6419

    lex.dur            lex.Cst        lex.Xst         dodth         sex           dobth
 Min.   :0.0001711   Alive :88613   Alive :87550   Min.   :2002   M:78001   Min.   :1906
 1st Qu.:0.2500000   DN    :45776   DN    :45498   1st Qu.:2006   F:56388   1st Qu.:1935
 Median :0.2500000   CVD   :    0   CVD   :  473   Median :2008             Median :1943
 Mean   :0.2449118   Cancer:    0   Cancer:  285   Mean   :2008             Mean   :1944
 3rd Qu.:0.2500000   Other :    0   Other :  583   3rd Qu.:2010             3rd Qu.:1952
 Max.   :0.2500000                                 Max.   :2011             Max.   :1990
                                                   NA's   :112191
      dodm           exit          entry           codth           docvd           cvd
 Min.   :1951   Min.   :2002   Min.   :2002          :112191   Min.   :1989   Min.   :1
```

```
1st Qu.:1990   1st Qu.:2012   1st Qu.:2002   CVD      :  7615   1st Qu.:1997    1st Qu.:1
Median :1996   Median :2012   Median :2002   Other    :  5195   Median :2001    Median :1
Mean   :1995   Mean   :2012   Mean   :2003   Cancer   :  4450   Mean   :2001    Mean   :1
3rd Qu.:2001   3rd Qu.:2012   3rd Qu.:2004   Lung     :  1549   3rd Qu.:2004    3rd Qu.:1
Max.   :2011   Max.   :2012   Max.   :2011   Infection:  1448   Max.   :2012    Max.   :1
                                             (Other)  :  1941   NA's   :22563   NA's   :22563
     dodr             dr            donef            nef            doneu            neu
Min.   :1991    Min.   :1       Min.   :1992    Min.   :1       Min.   :1991    Min.   :1
1st Qu.:1996    1st Qu.:1       1st Qu.:1997    1st Qu.:1       1st Qu.:1996    1st Qu.:1
Median :2001    Median :1       Median :2001    Median :1       Median :2001    Median :1
Mean   :2001    Mean   :1       Mean   :2001    Mean   :1       Mean   :2000    Mean   :1
3rd Qu.:2005    3rd Qu.:1       3rd Qu.:2006    3rd Qu.:1       3rd Qu.:2003    3rd Qu.:1
Max.   :2012    Max.   :1       Max.   :2012    Max.   :1       Max.   :2012    Max.   :1
NA's   :60335   NA's   :60335   NA's   :81511   NA's   :81511   NA's   :62368   NA's   :62368
    CoD          comp94
Alive :112191   None:126710
CVD   :  7615   Prev:  7679
Cancer:  4450
Other : 10133
```

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%
59.64134 67.63860 74.16838 79.98905 85.80698
```

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

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

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

```
     10%      50%      90%
2003.088 2006.719 2010.130
```

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

```
     10%      50%      90%
1978.863 1991.327 2001.567
```

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

```
                25%       50%       75%
 0.000000  9.254377 15.488044 21.669897
```

```
> 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%
 0.000000  5.167533 10.088817 15.108586 21.239940
```

```
> 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 2.922656 5.574264 8.848049
```

```
> 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 )
```

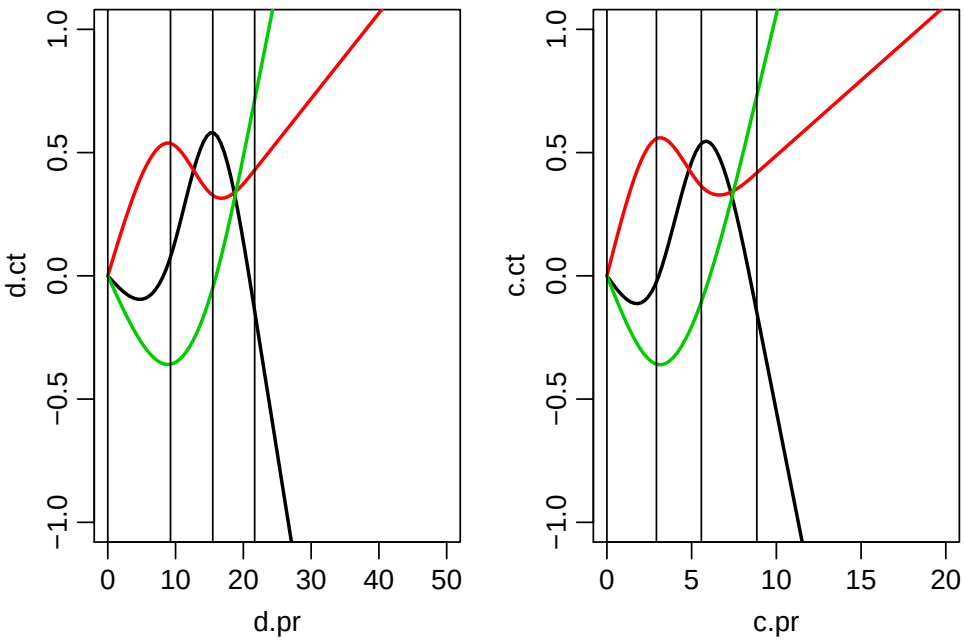

Figure 3.3: *The columns of the contrast matrices for diabetes and complications durations; note that for both duration codings, the value of all columns is 0 at 0.*

3.5 Including population mortality rates

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

```
> 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) )
```

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

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

```
Classes 'Lexis' and 'data.frame':       134384 obs. of  29 variables:
 $ sex    : Factor w/ 2 levels "M","F": 2 2 2 2 2 2 2 2 2 2 ...
 $ P      : num  2002 2002 2002 2002 2002 ...
 $ A      : num  17 17 17 17 17 17 17 17 17 17 ...
 $ lex.id : int  5844 1814 3378 5844 2222 5844 1814 5844 2222 2222 ...
 $ age    : num  17.3 17.2 17.9 17 17.8 ...
 $ Ddur   : num  4.147 1.975 0.037 3.897 0.27 ...
 $ per    : num  2002 2002 2002 2002 2003 ...
 $ DNdur  : num  0 0 0 0 0 0 0 0 0 0 ...
 $ lex.dur: num  0.25 0.25 0.25 0.25 0.25 ...
 $ lex.Cst: Factor w/ 5 levels "Alive","DN","CVD",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ lex.Xst: Factor w/ 5 levels "Alive","DN","CVD",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ dodth  : num  NA NA NA NA NA NA NA NA NA NA ...
 $ dobth  : num  1985 1985 1984 1985 1985 ...
 $ dodm   : num  1998 2000 2002 1998 2002 ...
 $ exit   : num  2012 2012 2012 2012 2012 ...
 $ entry  : num  2002 2002 2002 2002 2002 ...
 $ codth  : Factor w/ 10 levels "","CVD","Cancer",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ docvd  : num  NA NA 2002 NA NA ...
 $ cvd    : num  NA NA 1 NA NA NA NA NA NA NA ...
 $ dodr   : num  NA NA NA NA 2007 ...
 $ dr     : num  NA NA NA NA 1 NA NA NA 1 1 ...
 $ donef  : num  NA NA NA NA NA NA NA NA NA NA ...
 $ nef    : num  NA NA NA NA NA NA NA NA NA NA ...
 $ doneu  : num  NA NA NA NA NA NA NA NA NA NA ...
 $ neu    : num  NA NA NA NA NA NA NA NA NA NA ...
 $ CoD    : Factor w/ 4 levels "Alive","CVD",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ comp94 : Factor w/ 2 levels "None","Prev": 1 1 1 1 1 1 1 1 1 1 ...
 $ rate   : num  0.252 0.252 0.252 0.252 0.252 ...
```

```
 $ E      : num  6.3e-05 6.3e-05 6.3e-05 6.3e-05 6.3e-05 ...
 - 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 87546   450 205    159   248     88608     1062   21697.65      4185
  DN        0 45048 267    126   335     45776      728   11214.76      2109
  Sum   87546 45498 472    285   583    134384     1790   32912.41      5844
```

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 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)                  1.740 1.285  2.355
Ns(age, knots = a.kn)1       2.339 1.773  3.087
Ns(age, knots = a.kn)2       4.507 3.542  5.734
Ns(age, knots = a.kn)3      11.134 8.203 15.111
Ns(age, knots = a.kn)4       6.057 4.619  7.943
Ns(per, knots = p.kn)1       0.758 0.594  0.966
Ns(per, knots = p.kn)2       0.755 0.648  0.878
Ns(Ddur, knots = d.kn)1      1.803 1.001  3.246
Ns(Ddur, knots = d.kn)2      2.400 0.658  8.749
Ns(Ddur, knots = d.kn)3      2.263 1.050  4.877
I(age * Ddur/100)            0.956 0.899  1.016
I(lex.Cst == "DN")TRUE       1.862 1.657  2.091
comp94Prev                   1.128 0.939  1.354
```

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
  82.94    5.24   89.97
```

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

```
        cod  AllD         SMR
        sex     M     F     M     F
test
+AoD        0.275 0.993 0.260 0.988
-AoF        0.432 0.080 0.735 0.069
-AoD        0.000 0.000 0.000 0.000
+PoD        0.213 0.847 0.201 0.805
-PoF        0.919 0.115 0.973 0.081
-PoD        0.000 0.177 0.053 0.761
-comp94     0.174 0.612 0.180 0.578
```

The conclusion from these p-values is that either age at diagnosis or current age could be used. Likewise there is no possibility to choose between period of follow-up and period of diagnosis, but at least for men some period effect is needed. Finally that there is no evidence of difference between patients with complications present in 1994 and those getting complications later.

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

3.7 Estimation and extraction of the results

When we fit a model we extract the results at the same time, which in this case will be the age-specific rates for persons diagnosed in ages 25, 30,. . . ,75, and followed till age 90. However, even if we use current date in the model, we will make predictions ignoring this

by fixing the value for the date of follow-up to 1.1.2010, our reference point for the calendar time effect.

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

Thus we will extract three things from the model:

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

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

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

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

As a technical aside, note that the prediction frame we construct contains rows of `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(25,75,5) )
+ new.frame <- rbind( new.frame,
+                     make.frame(da) )
> str( new.frame )


'data.frame':        891 obs. of  7 variables:
 $ age    : num  NA 25.5 26 26.5 27 27.5 28 28.5 29 29.5 ...
 $ per    : num  NA 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...
 $ Ddur   : num  NA 0.5 1 1.5 2 2.5 3 3.5 4 4.5 ...
 $ lex.Cst: Factor w/ 3 levels "Alive","DN","Dead": NA 1 1 1 1 1 1 1 1 ...
 $ lex.dur: num  NA 1000 1000 1000 1000 1000 1000 1000 1000 1000 ...
 $ E      : num  NA 1 1 1 1 1 1 1 1 ...
 $ aD     : num  NA 25 25 25 25 25 25 25 25 25 ...


> # A prediction frame for persons with DN
> DN.frame <- transform( new.frame,
+                        lex.Cst=factor( as.integer(lex.Cst)+1,
+                                        levels=1:3,
+                                        labels=levels(lex.Cst) ) )
```

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

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

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

Since we want to give the RR between patients with complications and patients without both at 2010 and at 2002, we need a contrast matrix to apply to the parameters of the

model `apdcll`. If we name the three parameters from the terms `I(lex.Cst=="DN")` and `I(per-2010):I(lex.Cst=="DN")` as $(\gamma, \beta_0, \beta_1)$, 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") )
+ apD   <- update( apd  , . ~  Ns( age , knots=a.kn ) + I(age-Ddur)
+                           + Ns( per , knots=p.kn )
+                           + Ns( Ddur, knots=d.kn ) )
+ # Tests:
+ ComplTt[rs,sx,,] <- as.matrix( abs( anova( apdcs,
+                                            apdci,
+                                            apdcl,
+                                            apdcll,
+                                            apdc,
+                                            apdcp,
+                                            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 )
+ zn <- predict( apd  , newdata=new.frame, type="link", se.fit=TRUE )
+ AMort[rs,"Final","Alive",sx,,] <- exp( cbind(zf$fit,zf$se.fit) %*% ci.mat() )
+ AMort[rs,"Main" ,"Alive",sx,,] <- exp( cbind(zm$fit,zm$se.fit) %*% ci.mat() )
+ AMort[rs,"Int"  ,"Alive",sx,,] <- exp( cbind(zi$fit,zi$se.fit) %*% ci.mat() )
```

```
+ AMort[rs,"Naive","Alive",sx,,] <- exp( cbind(zn$fit,zn$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() )
+
+ # Age at diagnosis effects
+ AgeDM[rs,sx,] <- ci.exp( apD, subset="I\\(" )
+
+ # 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 )
+ CurveRR[rs,"Naive","Alive",sx,,] <- ci.exp( apd  , subset="per"     , ctr.mat=cbind(p.pr-p.rf) )
+
+ # 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
  82.91    1.15   88.98
```

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

```
                       exp(Est.)      2.5%     97.5%
(Intercept)            2.4143201 1.5196521 3.8357079
Ns(age, knots = a.kn)1 0.4555361 0.2976718 0.6971207
Ns(age, knots = a.kn)2 0.7772101 0.5889821 1.0255925
Ns(age, knots = a.kn)3 0.5411388 0.3979321 0.7358824
Ns(age, knots = a.kn)4 0.5717310 0.4464008 0.7322484
Ns(Ddur, knots = d.kn)1 0.9580779 0.6972619 1.3164540
Ns(Ddur, knots = d.kn)2 0.6658865 0.2576499 1.7209590
Ns(Ddur, knots = d.kn)3 1.0722389 0.8420628 1.3653330
I(lex.Cst == "DN")TRUE  1.9926924 1.6724267 2.3742882
I(per - 2010)           0.9862718 0.9540445 1.0195876
```

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

```
                                        exp(Est.)      2.5%     97.5%
(Intercept)                             2.4248983 1.5084862 3.8980349
Ns(age, knots = a.kn)1                  0.4597599 0.3004526 0.7035358
Ns(age, knots = a.kn)2                  0.7694033 0.5826289 1.0160523
Ns(age, knots = a.kn)3                  0.5367366 0.3946416 0.7299946
Ns(age, knots = a.kn)4                  0.5697035 0.4447888 0.7296994
Ns(Ddur, knots = d.kn)1                 0.9768536 0.7104202 1.3432091
Ns(Ddur, knots = d.kn)2                 0.6875832 0.2654226 1.7811994
```

```
Ns(Ddur, knots = d.kn)3                          1.0875588 0.8537523 1.3853948
I(lex.Cst == "DN")TRUE                           2.0108237 1.4328838 2.8218702
I(lex.Cst == "DN")FALSE:Ns(per, knots = p.kn)1 1.1987803 0.6944684 2.0693157
I(lex.Cst == "DN")TRUE:Ns(per, knots = p.kn)1  1.0430612 0.5983416 1.8183203
I(lex.Cst == "DN")FALSE:Ns(per, knots = p.kn)2 0.6521453 0.4636892 0.9171951
I(lex.Cst == "DN")TRUE:Ns(per, knots = p.kn)2  1.0309941 0.7323797 1.4513631
```

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

```
                        exp(Est.)      2.5%     97.5%
(Intercept)             2.4294405 1.5282849 3.8619640
Ns(age, knots = a.kn)1  0.4554513 0.2975923 0.6970473
Ns(age, knots = a.kn)2  0.7836084 0.5929695 1.0355374
Ns(age, knots = a.kn)3  0.5424703 0.3989120 0.7376915
Ns(age, knots = a.kn)4  0.5732549 0.4475306 0.7342988
Ns(Ddur, knots = d.kn)1 0.9549534 0.6946931 1.3127178
Ns(Ddur, knots = d.kn)2 0.6650127 0.2572458 1.7191411
Ns(Ddur, knots = d.kn)3 1.0582754 0.8270709 1.3541124
I(lex.Cst == "DN")TRUE  1.9592696 1.6267640 2.3597383
I(per - 2010)           0.9876584 0.9550217 1.0214105
comp94Prev              1.0938205 0.7960679 1.5029413
```

```
> ci.exp( apd )
```

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

```
> save( AMort, AgeDM, CurveRR, ComplRR, ComplTt, file="./data/T2AllCau.Rda" )
```

The average change in mortality by age at onset are:

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

```
              what   Est    lo    hi
cod       sex
All cause M        -3.22 -4.33 -2.10
          F        -1.54 -2.91 -0.15
SMR       M        -3.22 -4.33 -2.10
          F        -1.53 -2.90 -0.14
```

In order to assess how different these slopes by age at diagnosis are, we need a small function to compute the ratio of the rates from the confidence intervals:

```
> r2rr <-
+ function( r1, r2 )
+ {
+  l1 <- log( r1[1] )
+ sl1 <- log( r1[3]/r1[2] )/(2*1.96)
+  l2 <- log( r2[1] )
+ sl2 <- log( r2[3]/r2[2] )/(2*1.96)
+  lr <- l1-l2
+ slr <- sqrt( sl1^2 + sl2^2 )
+ res <- c( exp( c(lr,slr) %*% ci.mat() ),
+           1-pchisq( (lr/slr)^2, 1 ) )
+ names( res ) <- c("RR","lo","hi","P")
+ res
+ }
> round( ftable(AgeDM[1:2,,]-1)*100, 1 )
```

```
             what  Est   lo   hi
cod       sex
All cause M        -3.2 -4.3 -2.1
          F        -1.5 -2.9 -0.2
SMR       M        -3.2 -4.3 -2.1
          F        -1.5 -2.9 -0.1
```

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

```
   RR    lo    hi     P
0.983 0.965 1.001 0.062
```

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

```
   RR    lo    hi     P
0.983 0.965 1.001 0.061
```

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

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

```
> str( ComplTt)
```

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

```
> round( ftable( ComplTt[1:2,,,"Pval"], row.vars=3, col.vars=c(1,2) ), 3 )
```

```
                          cod All cause         SMR
                          sex         M     F     M     F
pred
Linear DdurxAge interaction         0.176 0.964 0.183 0.956
Linear CxP interaction              0.281 0.248 0.292 0.259
Linear period effect                0.962 0.120 0.931 0.086
No CxP interaction                  0.095 0.283 0.084 0.281
Compl 1994 ne Compl                 0.174 0.616 0.180 0.582
No Compl effect                     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           1.17 0.93 1.46 1.09 0.79 1.49 1.17 0.93 1.46 1.09 0.80 1.50
Compl later vs. None           1.71 1.47 1.99 1.97 1.63 2.37 1.71 1.47 1.99 1.96 1.63 2.36
Compl vs. None                 1.76 1.52 2.03 2.00 1.68 2.38 1.76 1.52 2.03 1.99 1.67 2.37
Change / year                  0.95 0.92 0.97 0.97 0.94 1.00 0.97 0.94 1.00 0.99 0.95 1.02
Compl vs. None (2002)          1.42 1.07 1.89 1.69 1.19 2.40 1.41 1.06 1.88 1.69 1.19 2.39
Compl vs. None (2010)          2.06 1.62 2.62 2.26 1.70 3.01 2.07 1.63 2.64 2.26 1.70 3.00
Change / year (None)           0.92 0.88 0.96 0.95 0.91 1.00 0.94 0.90 0.98 0.97 0.93 1.01
Change / year (Compl)          0.96 0.93 1.00 0.99 0.94 1.04 0.99 0.95 1.02 1.01 0.96 1.06
```

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

```
> ci.indep <-
+ function( EE )
+ {
+ # Assume that EE is a 2-row matrix with 3 columnas:
+ # Estimate, lower and upper ci
+ le <- log( EE[,1] )
+ sl <- log(EE[,3]/EE[,2])/(1.96*2)
+ dl <- diff(le)
+ sd <- sqrt(sum(sl^2))
+ res <- c( exp(dl),
+           exp(dl-1.96*sd),
+           exp(dl+1.96*sd),
+           1-pchisq((dl/sd)^2,1) )
```

```
+ names( res ) <- c("RR","lo","up","Pval")
+ res
+ }
> ftable( ComplRR[c("All cause","SMR"),,"Change / year",] )


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


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


   RR    lo    up  Pval
1.023 0.980 1.068 0.293


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


   RR    lo    up  Pval
1.017 0.975 1.062 0.433
```

We then transform this to the relevant numbers for the table in the paper, that is, RR at fixed times and % change for the mortality trends.

```
> 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] 1.757 1.756 1.998 1.993 0.945 ...
 - 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     1.8   1.5   2.0   2.0   1.7   2.4
                        SMR           1.8   1.5   2.0   2.0   1.7   2.4
Change / year           All cause    -5.5  -8.0  -2.9  -3.3  -6.4   0.0
                        SMR          -3.0  -5.6  -0.4  -1.4  -4.6   2.0
Compl vs. None (2002)   All cause     1.4   1.1   1.9   1.7   1.2   2.4
                        SMR           1.4   1.1   1.9   1.7   1.2   2.4
Compl vs. None (2010)   All cause     2.1   1.6   2.6   2.3   1.7   3.0
                        SMR           2.1   1.6   2.6   2.3   1.7   3.0
Change / year (None)    All cause    -8.1 -11.9  -4.1  -4.9  -9.2  -0.5
                        SMR          -5.8  -9.7  -1.7  -3.0  -7.4   1.5
Change / year (Compl)   All cause    -3.7  -7.0  -0.3  -1.4  -6.1   3.5
                        SMR          -1.1  -4.5   2.4   0.5  -4.2   5.5
```

From this we see that there is a significant decrease of 5.5% per year for male T2D and a 3.3% borderline significant decrease in 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(4,7,10) ## Age at dx 40 55 70
> mlim <- c(1,900)/3
> par( mar=c(3,3.5,1,1), mgp=c(3,1,0)/1.6 )
> plot( NA, type="n", log="y", ylim=mlim, xlim=c(40,90),
+       xlab="Age", ylab="", las=1 )
> abline( v=seq(0,100,5), h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
> for( sx in 2:1 ) # sx <- 1
+ matlines( pr.A[wh],
+           cbind( AMort["All cause","Final","Alive",sx,wh,],
+                  AMort["All cause","Final","DN",sx,wh,] ),
+           lwd=c(3,1,1), lty=rep(c(1,3),each=3), col=c("blue","red")[sx] )
> mtext( paste("Mortality at 1 Jan", pref, "(per 1000 PY)"), line=2.5, side=2, outer=FALSE )
> box()
```

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

```
> rlim <- c(1/2,10)
> par( mar=c(3,3.5,1,1), mgp=c(3,1,0)/1.6 )
> plot( NA, type="n", log="y", ylim=rlim, xlim=c(40,90),
+       xlab="Age", ylab="", las=1 )
> abline( v=seq(0,100,5), h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
> 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"]] )
> par( mar=c(3,3.5,1,1), mgp=c(3,1,0)/1.6 )
> plot( NA, type="n", log="y", ylim=rlim, xlim=c(2002,2011),
+       xlab="Date of FU", ylab="", las=1 )
> abline( v=2000:2015, h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
```

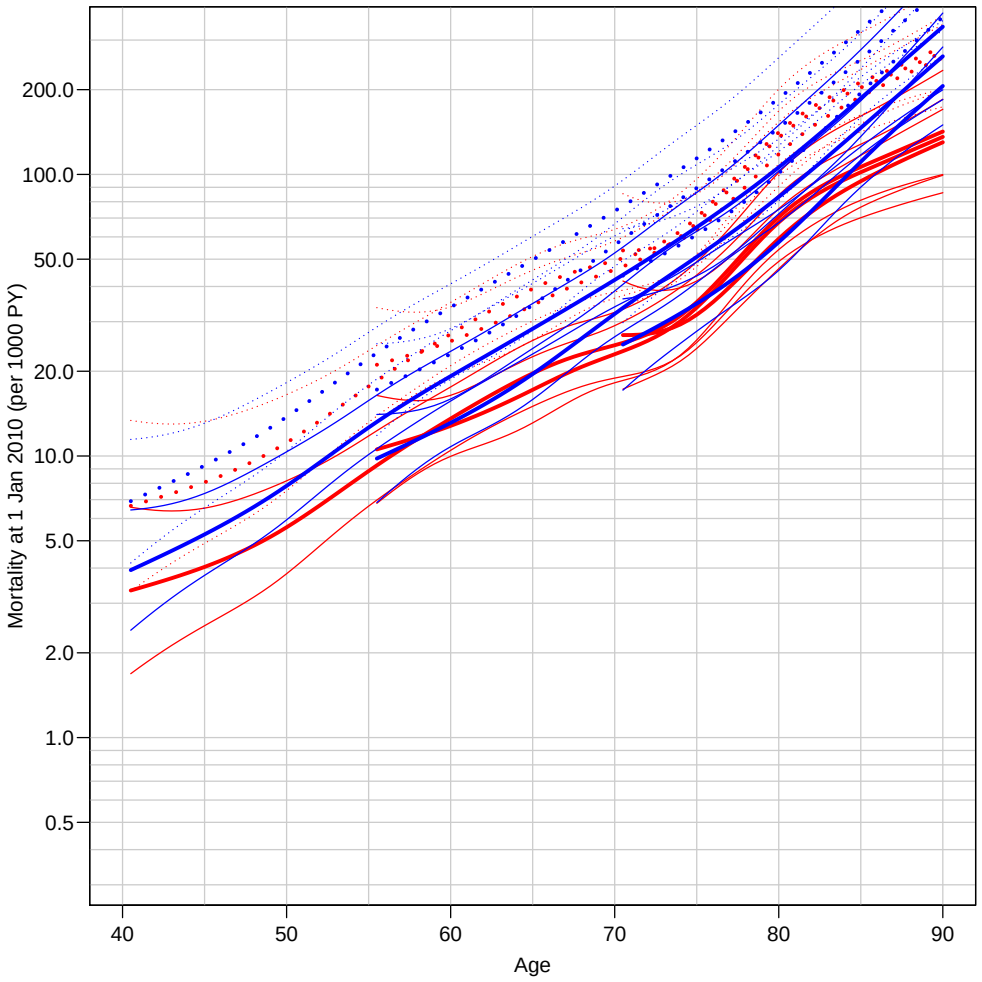

Figure 3.4: *Age-specific all-cause mortality rates without (full lines) and with (broken lines) DN for T2 patients aged 40, 55 and 75 at diagnosis (assumed to be in 2010). Red: F; blue: M; thin lines: 95% c.i.*

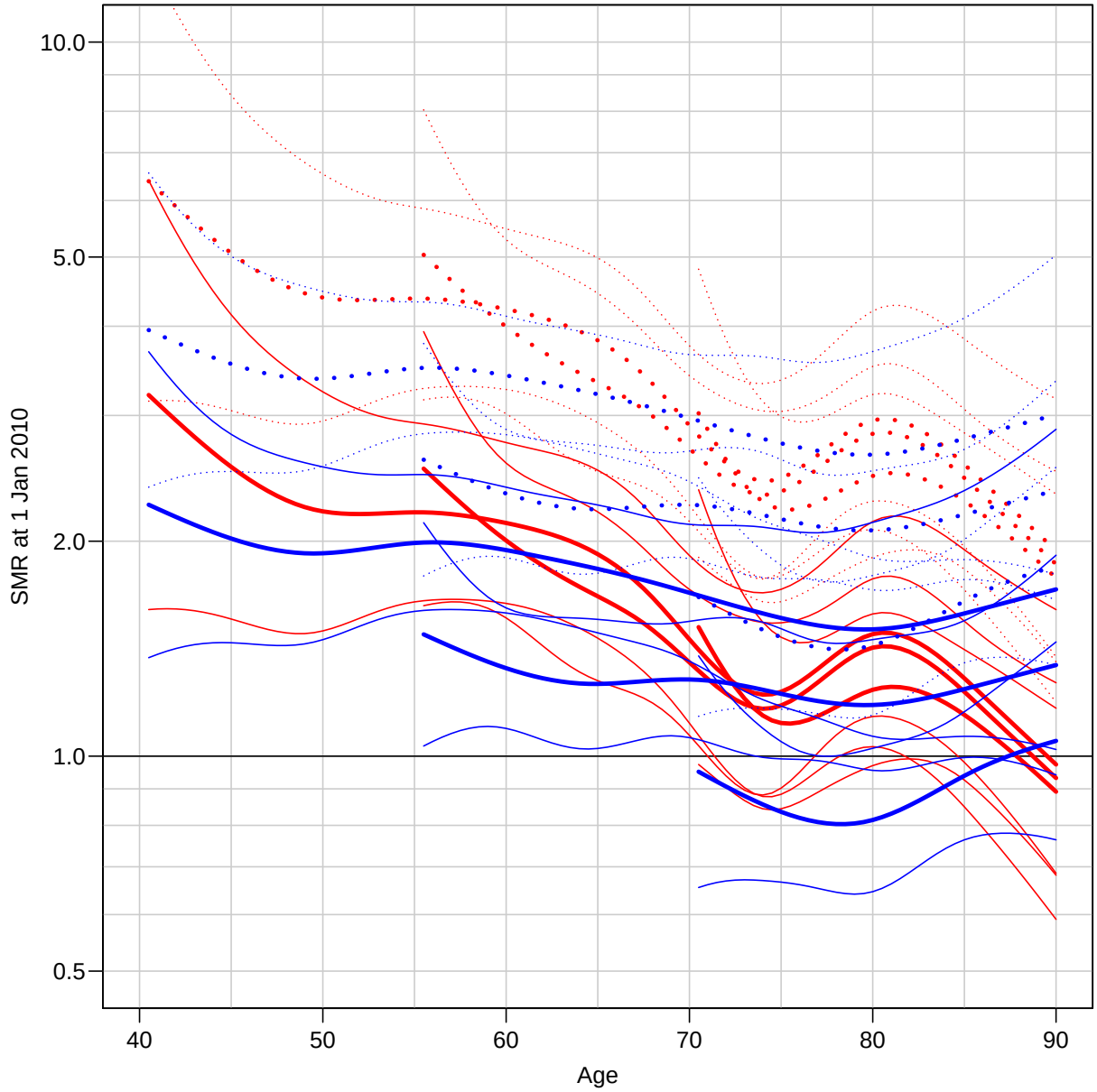

Figure 3.5: *Age-specific all-cause relative mortality (SMR). T2 patients aged 15, 30, 45 at diagnosis, T2 patients age 40, 55, 70 at diagnosis. Red: F; blue: M; thin lines: 95% c.i.*

```
> 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 for the two models together:

```
> par( mfrow=c(2,4), 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(40,90), xlab="Age", ylab="", las=1 )
+ abline( v=seq(0,100,5), h=outer(1:15,10^(-1:2),"*"), col=gray(0.8) )
+ if( rs=="All cause" )
+   mtext( paste( md, "model"), side=3, line=1 )
+ if( md=="Final" )
+   mtext( c( paste("Mortality at 1 Jan", pref, "(per 1000 PY)"),
+             paste("SMR 1 Jan", pref ) )[1+(rs=="SMR")],
+          line=3, side=2 )
+ if(rs=="SMR") abline(h=1)
+ for( sx in 1:2 )
+ matlines( as.numeric( dimnames(AMort)[["pred"]] ),
+           AMort[rs,md,"Alive",sx,,],
+           lwd=2, lty=c(1,0,0),
+           col=c("blue","red")[sx] )
+ box()
+    }
```

3.7.1 Graphs and tables for the paper

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

```
> str( ComplRR )
```

```
 num [1:5, 1:2, 1:8, 1:3] 1.17 1.17 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"
```

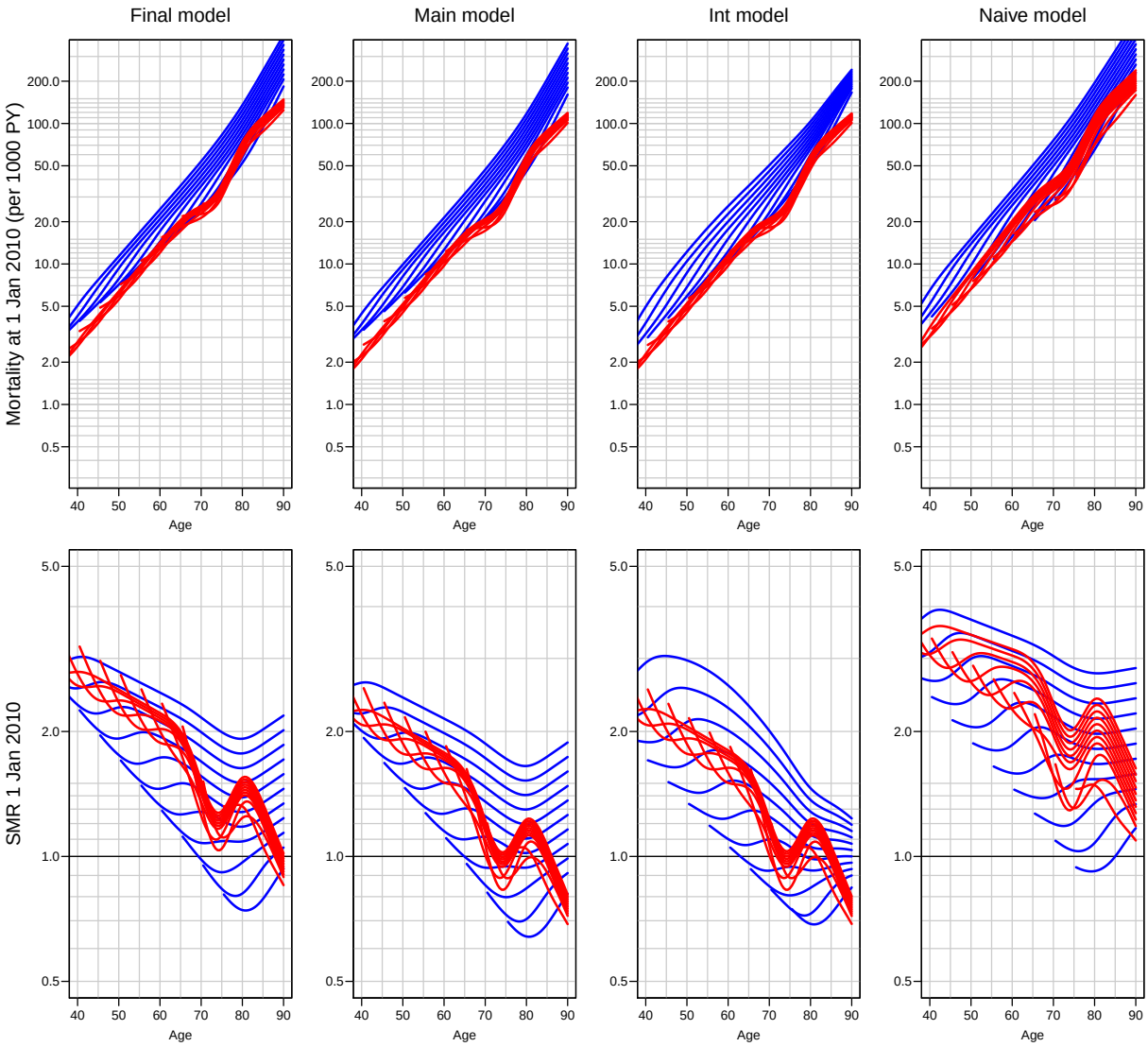

Figure 3.6: *Age-specific all-cause mortality and SMR for T2D patients without complications, diagnosed 1.1.2010 in ages 25,30,...75. The final model have a linear effect of calendar time, the main and the interaction model a 2-parameter-spline effect of calendar time; in all models the calendar time effect is separate for persons with and without complications. The interaction model has an extra 1-parameter (product) interaction between current age and current duration of diabetes. Red: F; blue: M*

```
> arttab[,,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		1.17	0.93	1.46	1.09	0.79	1.49
	SMR		1.17	0.93	1.46	1.09	0.80	1.50
Compl later vs. None	All cause		1.71	1.47	1.99	1.97	1.63	2.37
	SMR		1.71	1.47	1.99	1.96	1.63	2.36
Compl vs. None	All cause		1.76	1.52	2.03	2.00	1.68	2.38
	SMR		1.76	1.52	2.03	1.99	1.67	2.37
Change / year	All cause		-5.50	-8.00	-2.90	-3.30	-6.40	0.00
	SMR		-3.00	-5.60	-0.40	-1.40	-4.60	2.00
Compl vs. None (2002)	All cause		1.42	1.07	1.89	1.69	1.19	2.40
	SMR		1.41	1.06	1.88	1.69	1.19	2.39
Compl vs. None (2010)	All cause		2.06	1.62	2.62	2.26	1.70	3.01
	SMR		2.07	1.63	2.64	2.26	1.70	3.00
Change / year (None)	All cause		-8.10	-11.90	-4.10	-4.90	-9.20	-0.50
	SMR		-5.80	-9.70	-1.70	-3.00	-7.40	1.50
Change / year (Compl)	All cause		-3.70	-7.00	-0.30	-1.40	-6.10	3.50
	SMR		-1.10	-4.50	2.40	0.50	-4.20	5.50

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

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

```
> f1 <-
+ function(cls=1:3,dr=TRUE,my=2.5,ry=0.55,aod=c(4,7,10))
+ {
+ 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% aod)               ## Age at dx indicator
+ fs <- which(diff(c(0,wh*agr))>0)+1 ## First point of predictions
+
+ p.pr <- as.numeric( dimnames(CurveRR)[["pred"]] )
+
+ ######################################################################
+ # Mortality
+ mlim <- c(2,500)
+ for( sx in 1:2 )
+    {
+ plot( NA, type="n", log="y", ylim=mlim, xlim=c(40,90),
+       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(40,90),
+       xlab="Age", ylab="", yaxt="n", las=1 )
+ abline( v=seq(0,100,5), h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
+ abline( h=1 )
+ matlines( pr.A[wh], AMort["SMR","Final","Alive",sx,wh,cls],
+           lwd=c(3,1,1), lty=1, col=scol[sx] )
+ matlines( pr.A[wh], AMort["SMR","Final","DN"   ,sx,wh,cls],
+           lwd=c(3,1,1), lty="11", lend=1, col=scol[sx] )
+ if( sx==1 )
+   {
+   axis(side=2)
+   mtext( paste("SMR at 1 Jan", pref ),
+          line=2.5, side=2, outer=FALSE, las=0, cex=0.7 )
+   }
+ if( !is.null(ry) )
+   {
+   segments( floor(pr.A[fs]),
+             pmin(ry,apply(AMort["SMR","Final",c("Alive","DN"),sx,fs,1],2,min)),
+             floor(pr.A[fs]),
+             pmax(ry,apply(AMort["SMR","Final",c("Alive","DN"),sx,fs,1],2,max)),
+             col=scol[sx] )
+   text( floor(pr.A[fs])+1, ry, paste(floor(pr.A[fs])), adj=0, col=scol[sx] )
+   }
+ box()
+ mtext( c("c","d")[sx], side=3, adj=0.01, line=0.2 )
+   }
+
+ if( dr ) {
+ ##############################################################################
+ # Mortality RR
+ rlim <- c(0.4,4)
+ for( sx in 1:2 )
+    {
+ plot( NA, type="n", log="y", ylim=rlim, xlim=c(2002,2011),
+       xlab="Date of follow-up", ylab="", las=1, yaxt="n" )
+ abline( v=2000:2015, h=outer(1:9,10^(-1:2),"*"), col=gray(0.8) )
+ abline( h=1 )
+ matlines( p.pr, CurveRR["All cause","Final","Alive",sx,,cls],
+           lwd=c(3,1,1), lty=1, col=scol[sx] )
+ matlines( p.pr, CurveRR["SMR"      ,"Final","Alive",sx,,cls],
+           lwd=c(3,1,1), lty="11", lend=1, col=scol[sx] )
+ if( sx==1 )
+   {
+   axis(side=2)
+   mtext( "SMR ratio (broken)     Mortality ratio (full)",
+          line=2.5, side=2, outer=FALSE, las=0, cex=0.7 )
+   }
+ box()
+ mtext( c("e","f")[sx], side=3, adj=0.01, line=0.2 )
+   } }
+ }
> f1(aod=c(4,6,8,10))
```

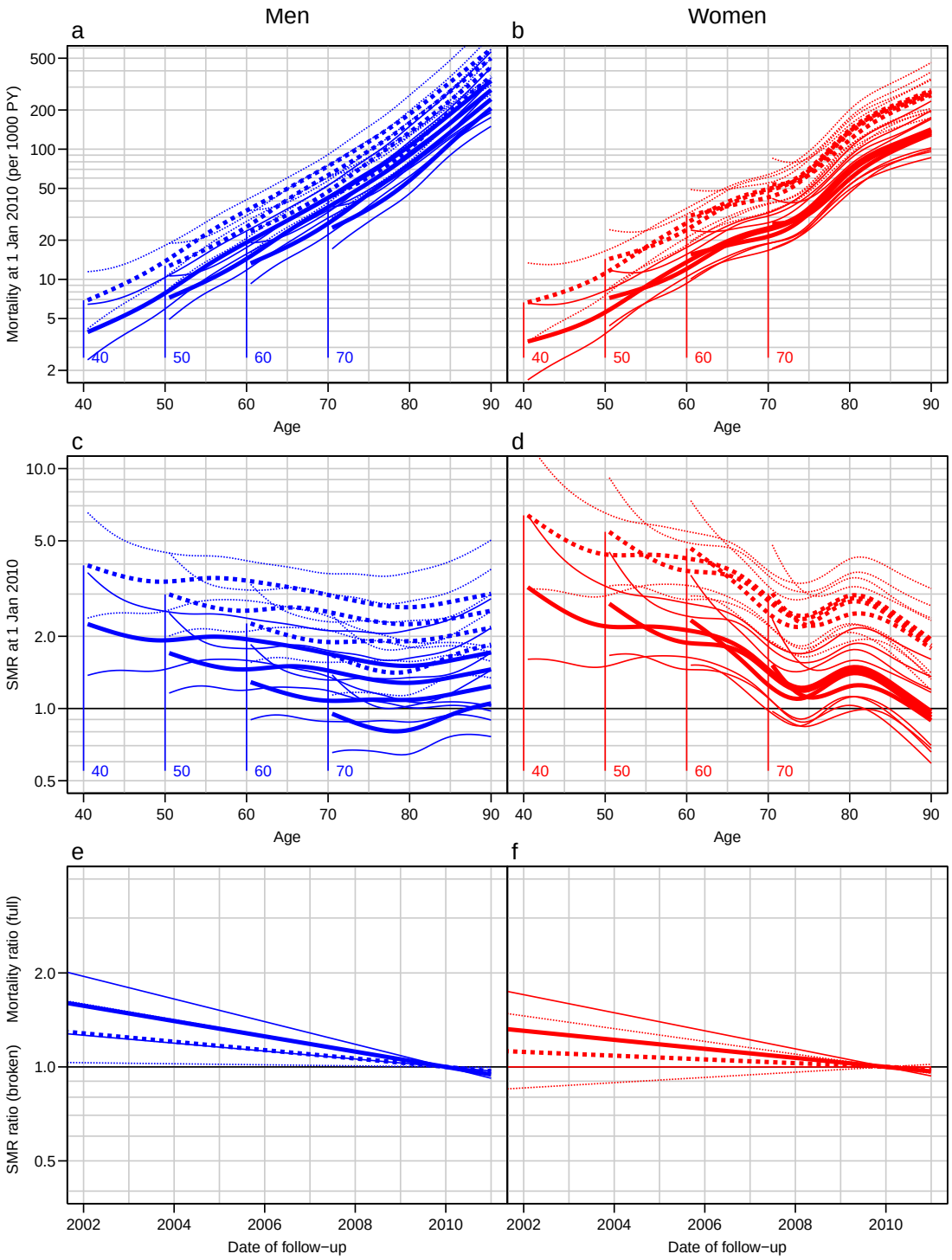

Figure 3.7: *Mortality rates (a,b), SMR (c,d) and changes in these (e,f). Patients diagnosed with DM in ages 40, 50, 60 and 70. For mortality rates and SMR, patients without nephropathy are with full lines, patients with nephropathy are with dotted lines. For the changes (e,f), the full lines are changes in mortality, dotted lines changes in SMR. Thin lines indicate 95% confidence intervals throughout.*

```
> f1( cls=1,aod=4:10 )
> pdf( "T2Fig1.pdf", height=9, width=7 )
> f1( cls=1, aod=c(4,6,8,10) )
> dev.off()


pdf
  2


> win.metafile( "T2Fig1.emf", height=9, width=7 )
> f1( cls=1, aod=c(4,6,8,10) )
> dev.off()


pdf
  2
```

3.7.2 Results from the naive model

We also show results for the overall mortality and SMR as estimated from the naïve model where nephropathy status is ignored. Hence we only have one set of curves in each graph, referring to the overall mortality regardless of nepropathuy status.

```
> f1n <-
+ function(cls=1:3,dr=TRUE,my=2.5,ry=0.55,aod=c(4,7,10))
+ {
+ 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% aod)                ## Age at dx indicator
+ fs <- which(diff(c(0,wh*agr))>0)+1 ## First point of predictions
+
+ p.pr <- as.numeric( dimnames(CurveRR)[["pred"]] )
+
+ ##########################################################################
+ # Mortality
+ mlim <- c(2,500)
+ for( sx in 1:2 )
+    {
+ plot( NA, type="n", log="y", ylim=mlim, xlim=c(40,90),
+       xlab="Age", ylab="", las=1, yaxt="n" )
+ abline( v=seq(0,100,5), h=outer(c(1.5,1:9),10^(-1:2),"*"), col=gray(0.8) )
+ matlines( pr.A[wh], AMort["All cause","Naive","Alive",sx,wh,cls],
+           lwd=c(3,1,1), lty=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,AMort["All cause","Naive","Alive",sx,fs,1]),
+             floor(pr.A[fs]),
+             pmax(my,AMort["All cause","Naive","Alive",sx,fs,1]),
+             col=scol[sx] )
+   text( floor(pr.A[fs])+1, my, paste(floor(pr.A[fs])), adj=0, col=scol[sx] )
+   }
```

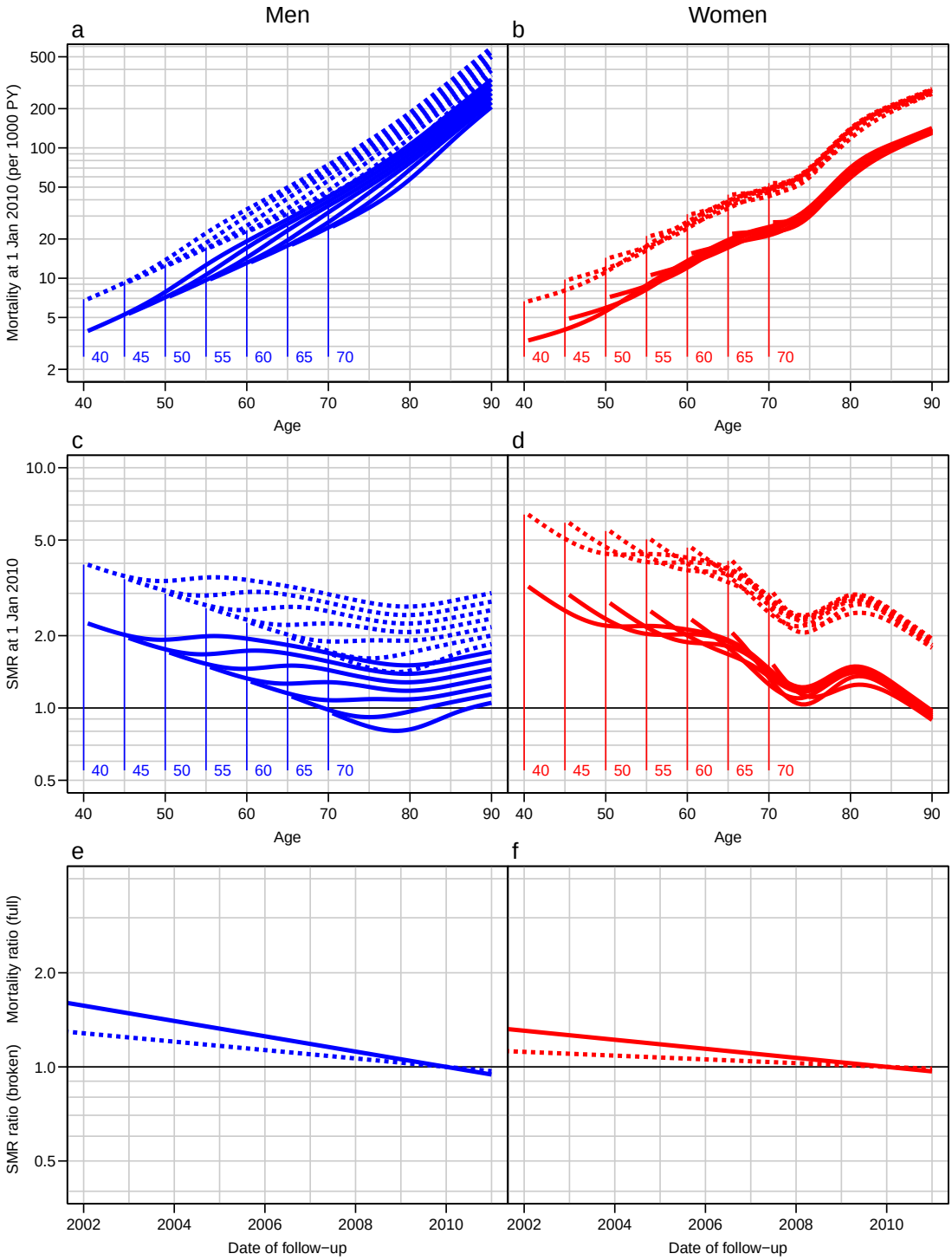

Figure 3.8: *Mortality rates (a,b), SMR (c,d) and changes in these (e,f). Patients diagnosed with DM in ages 40, 45,...,70. For mortality rates and SMR, patients without nephropathy are with full lines, patients with nephropathy are with dotted lines. For the changes (e,f), the full lines are changes in mortality, dotted lines changes in SMR.*

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

```
pdf
  2
```

```
> win.metafile( "T2Fig1n.emf", height=9, width=7 )
> f1n( cls=1, aod=c(4,6,8,10) )
> dev.off()
```

```
pdf
  2
```

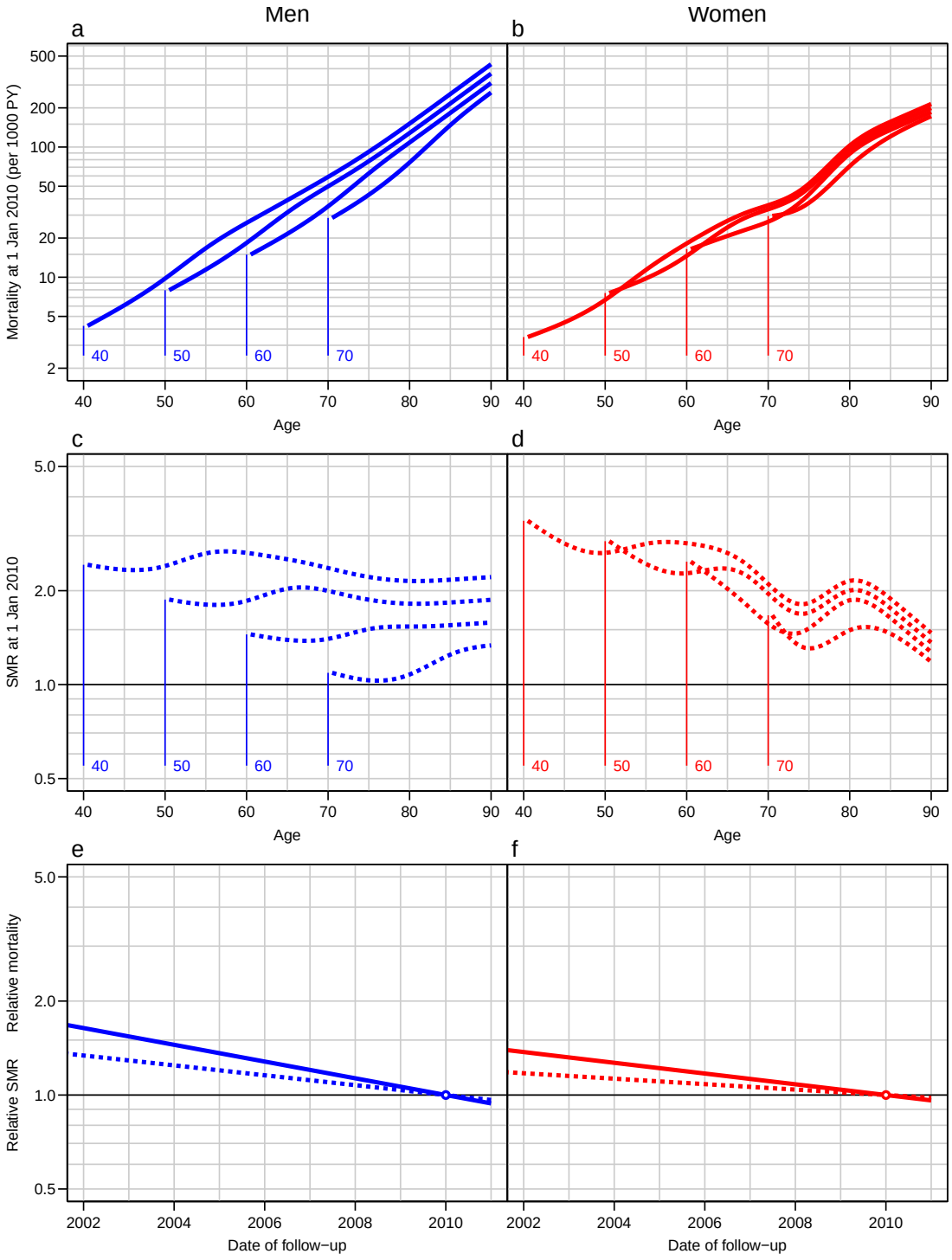

Figure 3.9: *Overall mortality rates (a,b), SMR (c,d) and changes in these (e,f). Patients diagnosed with DM in ages 40, 50, 60 and 70. Mortality rates are full lines, SMRs are with broken lines.*