Trends in prevalence, incidence and mortality of T1 and T2 diabetes in Denmark 1996–2017

Electronic Supplementary Material

SDCC
March 2018
http://bendixcarstensen.com/
Version 1

Compiled Wednesday 4th April, 2018, 18:42
from: `/home/bendix/sdc/DMreg/NewReg/DMDK/ESM/ESMx.tex`

Bendix Carstensen
Senior Statistician

Steno Diabetes Center Copenhagen, Gentofte, Denmark
Clinical Epidemiology
& Department of Biostatistics, University of Copenhagen
bcar0029@regionh.dk   b@bxc.dk
http://BendixCarstensen.com

Contents

. . . now input from `init.tex`

Chapter 1

Prevalence and ocurrence rates

1.1 Prerequisites

We will make models for the rates as observed in small intervals of age and calendar time, so we start by specifying the interval length, and then the points at which we want to predict. The transition rates are labeled by the midpoints of the Lexis squares (of width `int`) where we predict them (`a.pt` and `p.pt`), and the prevalences by the midpoints of the age-classes (`a.pt` and the time-points `t.pt` — note that we shall make predictions of rates all the way to 2030 — this will mainly be with the purpose of predicting the total number of DM persons in the population:

```
> int <- 1/12
> a.pt <- seq(0,100,int)[-1] - int/2
> t.pt <- seq(1996,2040,int)
> p.pt <- t.pt[-1] - int/2
```

We shall model all the rates by age-period-cohort models separately for men and women, both on a multiplicative and an additive scale. As a sensitivity analysis we will also model the rates only by an age-period model.

We will use natural splines to model the effects of age, period and cohort, and for all analyses we will use the same *number* of knots for these three effects, but of course place them differently based on the location of information, *i.e.* the events:

```
> nk.a <- 8
> nk.p <- 6
> nk.c <- 8
```

For the practical location of the spline knots we also define a small function which from the number of knots derives reasonable quantiles:

```
> qn <- function( nk, bd=2 ) seq( from =   1/(bd*nk),
+                                   to = 1-1/(bd*nk),
+                               length = nk )
> qn( 10, 1 )
 [1] 0.1000000 0.1888889 0.2777778 0.3666667 0.4555556 0.5444444 0.6333333 0.7222222
 [9] 0.8111111 0.9000000
> qn( 10, 2 )
 [1] 0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
> qn( 10, 4 )
```

```
 [1] 0.0250000 0.1305556 0.2361111 0.3416667 0.4472222 0.5527778 0.6583333 0.7638889
 [9] 0.8694444 0.9750000
```

So this function generates `nk`, equidistant points in the interval (0,1) where the outer points are 1/(2 × `nk`) from the end. The second parameter, `bd`, is for modifying the multiplier 2, so that the outer intervals are 1/(`nk` × `bd`):

Finally, we define a couple of functions to enhance readabilty of lareg numbers (> 5 digits):

```
> fC <- function( x, d=0, w=9, z=NULL ) formatC( x,
+                                           format = "f",
+                                         big.mark = ",",
+                                           digits = d,
+                                            width = w,
+                                       zero.print = z )
> fCp     <- function( x, d=0, w=9, z=".", ... ) noquote( fC( x, d=d, w=w, z=z ), ... )
> fCtable <- function( x, d=0, w=9, z=".", ... )  ftable( fC( x, d=d, w=w, z=z ), ... )
```

```
> save( qn, fC, fCp, fCtable,
+       int, a.pt, t.pt, p.pt, nk.a, nk.p, nk.c ,
+       file="../data/inits.Rda" )
```

. . . now input from `prev.tex`

Chapter 2

Diabetes prevalence

Prevalence data as of 1 January each year 1996–2017 are available from the prevalence tabulation of the diabetes register in the SAS-file `prv`:

```
> prv  <- read_sas("../data/prv.sas7bdat")
> str( prv )
Classes 'tbl_df', 'tbl' and 'data.frame':        63844 obs. of  7 variables:
 $ pdat : num  1995 1995 1995 1995 1995 ...
 $ REG  : atomic  81 81 81 81 81 81 81 81 81 81 ...
  ..- attr(*, "label")= chr "Region i Danmark"
 $ state: chr  "T1" "T1" "T1" "T1" ...
 $ sex  : num  1 1 1 1 1 1 1 1 1 1 ...
 $ age  : num  2 3 5 6 7 8 9 10 11 12 ...
 $ n    : num  2 2 3 2 2 2 7 7 4 8 ...
 $ a5   : num  0 0 5 5 5 5 5 10 10 10 ...
 - attr(*, "label")= chr "PRV"
> prv$sex <- factor( prv$sex, labels=c("M","F") )
```

2.0.1 Raw tables

Here we produce the prevalence tables needed in the paper:

```
> tt <- xtabs( n ~ pdat + sex + state, data=prv )
> fCtable( mm <- addmargins( tt[,,-3], margin=2:3 ), col.vars=3:2 )
     state        T1                              T2                              Sum          
     sex           M         F       Sum           M         F       Sum           M         F
pdat                                                                                           
1995          12,204     9,664    21,868      22,573    24,205    46,778      34,777    33,869
1996          13,287    10,365    23,652      34,938    35,037    69,975      48,225    45,402
1997          13,698    10,610    24,308      38,274    37,713    75,987      51,972    48,323
1998          13,970    10,871    24,841      41,585    40,178    81,763      55,555    51,049
1999          14,278    11,031    25,309      45,352    42,964    88,316      59,630    53,995
2000          14,519    11,211    25,730      49,337    46,100    95,437      63,856    57,311
2001          14,790    11,360    26,150      53,134    49,023   102,157      67,924    60,383
2002          15,044    11,460    26,504      56,747    51,778   108,525      71,791    63,238
2003          15,183    11,592    26,775      61,359    55,882   117,241      76,542    67,474
2004          15,343    11,663    27,006      66,768    60,534   127,302      82,111    72,197
2005          15,510    11,788    27,298      72,389    65,209   137,598      87,899    76,997
2006          15,635    11,954    27,589      76,960    68,709   145,669      92,595    80,663
2007          15,806    12,067    27,873      81,381    71,670   153,051      97,187    83,737
2008          15,965    12,206    28,171      86,302    75,387   161,689     102,267    87,593
```

```
2009          16,158    12,359    28,517    92,099    79,781   171,880   108,257    92,140
2010          16,328    12,468    28,796    98,131    84,128   182,259   114,459    96,596
2011          16,477    12,580    29,057   105,082    89,040   194,122   121,559   101,620
2012          16,595    12,722    29,317   115,294    98,025   213,319   131,889   110,747
2013          16,718    12,816    29,534   122,406   103,905   226,311   139,124   116,721
2014          16,809    12,948    29,757   126,318   106,913   233,231   143,127   119,861
2015          16,893    13,083    29,976   129,534   109,134   238,668   146,427   122,217
2016          17,024    13,220    30,244   133,784   112,095   245,879   150,808   125,315
```

```
> str( mm )
```

```
 table [1:22, 1:3, 1:3] 12204 13287 13698 13970 14278 ...
 - attr(*, "dimnames")=List of 3
  ..$ pdat : chr [1:22] "1995" "1996" "1997" "1998" ...
  ..$ sex  : chr [1:3] "M" "F" "Sum"
  ..$ state: chr [1:3] "T1" "T2" "Sum"
```

```
> # T1 as percentage of all DM
> fCtable( mm[,,1]/mm[,,3]*100, d=1 )
```

```
     sex         M         F       Sum
pdat
1995          35.1      28.5      31.9
1996          27.6      22.8      25.3
1997          26.4      22.0      24.2
1998          25.1      21.3      23.3
1999          23.9      20.4      22.3
2000          22.7      19.6      21.2
2001          21.8      18.8      20.4
2002          21.0      18.1      19.6
2003          19.8      17.2      18.6
2004          18.7      16.2      17.5
2005          17.6      15.3      16.6
2006          16.9      14.8      15.9
2007          16.3      14.4      15.4
2008          15.6      13.9      14.8
2009          14.9      13.4      14.2
2010          14.3      12.9      13.6
2011          13.6      12.4      13.0
2012          12.6      11.5      12.1
2013          12.0      11.0      11.5
2014          11.7      10.8      11.3
2015          11.5      10.7      11.2
2016          11.3      10.5      11.0
```

```
> # Crude prevalences
> xx <- addmargins( tt, margin=2:3 )
> str( xx )
```

```
 table [1:22, 1:3, 1:4] 12204 13287 13698 13970 14278 ...
 - attr(*, "dimnames")=List of 3
  ..$ pdat : chr [1:22] "1995" "1996" "1997" "1998" ...
  ..$ sex  : chr [1:3] "M" "F" "Sum"
  ..$ state: chr [1:4] "T1" "T2" "Well" "Sum"
```

```
> fCtable( mm/xx[,,c(4,4,4)]*100, col.vars=c(3:2), d=1 )
```

```
     state        T1                            T2                           Sum
     sex           M         F       Sum         M         F       Sum         M         F
pdat
1995             0.5       0.4       0.4       0.9       0.9       0.9       1.3       1.3
1996             0.5       0.4       0.5       1.3       1.3       1.3       1.9       1.7
```

```
1997             0.5       0.4       0.5       1.5       1.4       1.4       2.0       1.8
1998             0.5       0.4       0.5       1.6       1.5       1.5       2.1       1.9
1999             0.5       0.4       0.5       1.7       1.6       1.7       2.3       2.0
2000             0.6       0.4       0.5       1.9       1.7       1.8       2.4       2.1
2001             0.6       0.4       0.5       2.0       1.8       1.9       2.6       2.2
2002             0.6       0.4       0.5       2.1       1.9       2.0       2.7       2.3
2003             0.6       0.4       0.5       2.3       2.1       2.2       2.9       2.5
2004             0.6       0.4       0.5       2.5       2.2       2.4       3.1       2.6
2005             0.6       0.4       0.5       2.7       2.4       2.5       3.3       2.8
2006             0.6       0.4       0.5       2.9       2.5       2.7       3.4       2.9
2007             0.6       0.4       0.5       3.0       2.6       2.8       3.6       3.0
2008             0.6       0.4       0.5       3.2       2.7       3.0       3.8       3.2
2009             0.6       0.4       0.5       3.4       2.9       3.1       4.0       3.3
2010             0.6       0.4       0.5       3.6       3.0       3.3       4.2       3.5
2011             0.6       0.4       0.5       3.8       3.2       3.5       4.4       3.6
2012             0.6       0.5       0.5       4.1       3.5       3.8       4.7       3.9
2013             0.6       0.5       0.5       4.4       3.7       4.0       5.0       4.1
2014             0.6       0.5       0.5       4.5       3.7       4.1       5.1       4.2
2015             0.6       0.5       0.5       4.6       3.8       4.2       5.2       4.3
2016             0.6       0.5       0.5       4.7       3.9       4.3       5.2       4.3
```

2.0.2 Analysis dataset

For analysis we need to turn the dataset side-ways, so that the state-variable becomes three variables:

```
> pt1 <- subset( prv, state=="T1"  , select=c(1,2,4,5,6) ) ; names(pt1)[5]<-"T1"
> pt2 <- subset( prv, state=="T2"  , select=c(1,2,4,5,6) ) ; names(pt2)[5]<-"T2"
> pnd <- subset( prv, state=="Well", select=c(1,2,4,5,6) ) ; names(pnd)[5]<-"nD"
> prv <- merge( pt1, merge( pt2, pnd, all=TRUE ), all=TRUE )
> prv[is.na(prv)] <- 0
> names( prv )[1:4] <- c("P","reg","sex","A")
> prv <- subset( prv, P>1995.5 & A<99.5 )
> prv <- transform( prv, N = T1+T2+nD,
+                       DM = T1+T2 )
> str( prv )
'data.frame':        21000 obs. of  9 variables:
 $ P  : num  1996 1996 1996 1996 1996 ...
 $ reg: num  81 81 81 81 81 81 81 81 81 81 ...
 $ sex: Factor w/ 2 levels "M","F": 1 1 1 1 1 1 1 1 1 1 ...
 $ A  : num  0 1 2 3 4 5 6 7 8 9 ...
 $ T1 : num  0 0 3 4 2 1 5 5 3 3 ...
 $ T2 : num  0 0 0 0 0 0 0 0 0 1 ...
 $ nD : num  3818 3931 3924 3863 3689 ...
 $ N  : num  3818 3931 3927 3867 3691 ...
 $ DM : num  0 0 3 4 2 1 5 5 3 4 ...
```

```
> fCtable( addmargins(
+          xtabs( cbind(T1,T2,DM,nD,N) ~ P + sex, data=prv ),
+                      margin=2 ),
+          row.vars=2:1 )
                 T1        T2        DM        nD         N
sex P
M   1996     13,287     34,935     48,222 2,545,711 2,593,933
    1997     13,698     38,272     51,970 2,553,868 2,605,838
    1998     13,970     41,582     55,552 2,560,437 2,615,989
```

	1999	14,278	45,350	59,628	2,565,749	2,625,377
	2000	14,519	49,333	63,852	2,569,924	2,633,776
	2001	14,790	53,128	67,918	2,576,441	2,644,359
	2002	15,044	56,741	71,785	2,582,666	2,654,451
	2003	15,183	61,352	76,535	2,586,727	2,663,262
	2004	15,343	66,763	82,106	2,588,532	2,670,638
	2005	15,508	72,383	87,891	2,590,231	2,678,122
	2006	15,635	76,954	92,589	2,594,694	2,687,283
	2007	15,805	81,376	97,181	2,601,746	2,698,927
	2008	15,964	86,298	102,262	2,613,689	2,715,951
	2009	16,157	92,095	108,252	2,628,640	2,736,892
	2010	16,327	98,126	114,453	2,637,488	2,751,941
	2011	16,477	105,076	121,553	2,646,540	2,768,093
	2012	16,595	115,285	131,880	2,651,497	2,783,377
	2013	16,717	122,395	139,112	2,660,040	2,799,152
	2014	16,809	126,308	143,117	2,674,186	2,817,303
	2015	16,893	129,522	146,415	2,695,379	2,841,794
	2016	17,024	133,767	150,791	2,722,296	2,873,087
F	1996	10,365	35,023	45,388	2,613,615	2,659,003
	1997	10,610	37,702	48,312	2,621,376	2,669,688
	1998	10,871	40,167	51,038	2,627,206	2,678,244
	1999	11,030	42,945	53,975	2,632,772	2,686,747
	2000	11,211	46,085	57,296	2,636,780	2,694,076
	2001	11,360	49,008	60,368	2,642,956	2,703,324
	2002	11,460	51,756	63,216	2,649,116	2,712,332
	2003	11,590	55,860	67,450	2,651,836	2,719,286
	2004	11,661	60,505	72,166	2,653,347	2,725,513
	2005	11,787	65,185	76,972	2,655,359	2,732,331
	2006	11,953	68,677	80,630	2,659,705	2,740,335
	2007	12,067	71,631	83,698	2,665,982	2,749,680
	2008	12,206	75,351	87,557	2,675,849	2,763,406
	2009	12,359	79,741	92,100	2,688,997	2,781,097
	2010	12,468	84,090	96,558	2,699,299	2,795,857
	2011	12,578	89,004	101,582	2,709,024	2,810,606
	2012	12,720	97,986	110,706	2,713,269	2,823,975
	2013	12,815	103,857	116,672	2,720,447	2,837,119
	2014	12,947	106,856	119,803	2,732,334	2,852,137
	2015	13,082	109,068	122,150	2,748,345	2,870,495
	2016	13,218	112,021	125,239	2,769,132	2,894,371
Sum	1996	23,652	69,958	93,610	5,159,326	5,252,936
	1997	24,308	75,974	100,282	5,175,244	5,275,526
	1998	24,841	81,749	106,590	5,187,643	5,294,233
	1999	25,308	88,295	113,603	5,198,521	5,312,124
	2000	25,730	95,418	121,148	5,206,704	5,327,852
	2001	26,150	102,136	128,286	5,219,397	5,347,683
	2002	26,504	108,497	135,001	5,231,782	5,366,783
	2003	26,773	117,212	143,985	5,238,563	5,382,548
	2004	27,004	127,268	154,272	5,241,879	5,396,151
	2005	27,295	137,568	164,863	5,245,590	5,410,453
	2006	27,588	145,631	173,219	5,254,399	5,427,618
	2007	27,872	153,007	180,879	5,267,728	5,448,607
	2008	28,170	161,649	189,819	5,289,538	5,479,357
	2009	28,516	171,836	200,352	5,317,637	5,517,989
	2010	28,795	182,216	211,011	5,336,787	5,547,798
	2011	29,055	194,080	223,135	5,355,564	5,578,699
	2012	29,315	213,271	242,586	5,364,766	5,607,352
	2013	29,532	226,252	255,784	5,380,487	5,636,271

```
    2014      29,756   233,164   262,920 5,406,520 5,669,440
    2015      29,975   238,590   268,565 5,443,724 5,712,289
    2016      30,242   245,788   276,030 5,491,428 5,767,458
```

We now have the data in the format we will use for analysis of the rates.

2.1 Smoothed prevalences

We model the prevalences as of 1 January each of the years 1996—2016, as a smooth function of age, and use the predicted prevalences to produce the prevalence of diabetes in each of the smaller age-classes that we use for the simulation. We use a log-link binomial model with a smooth spline with 12 knots, using different kots for T1, T2 and DM:

```
> qn <- function(n) (1:n-0.5)/n
>   kn1 <- c( 10, 20, 40, 60 )
> ( kn2 <- c( 15, with( prv, quantile( rep(A,T2), qn(10) ) ) ) )
     5% 15% 25% 35% 45% 55% 65% 75% 85% 95% 
 15  42  51  56  60  64  67  71  75  79  86 
> ( knd <- c( 15, with( prv, quantile( rep(A,DM), qn(10) ) ) ) )
     5% 15% 25% 35% 45% 55% 65% 75% 85% 95% 
 15  32  46  53  58  62  66  69  73  78  85 
```

We fit a model for the prevalence of T1 and T2 and T1+T2 (DM) for each date of interest, and store the resulting fitted age-specific prevelences in an array, we set up:

```
> a.pt <- seq(0,100,0.5)
> t.pt <- 1996:2017
> parr <- NArray( list( typ = c("T1","T2","DM"),
+                       sex = levels( prv$sex ),
+                         A = a.pt,
+                         T = t.pt,
+                             c("Est","lo","hi") ) )
> tarr <- NArray( dimnames(parr)[c(1:2,5)] )
> for( tp in dimnames(parr)[["typ"]] )
+    {
+    prv$X <- prv[,tp]
+    a.kn <- switch( tp, T1 = kn1,
+                        T2 = kn2,
+                        DM = knd )
+ for( sx in dimnames(parr)[["sex"]] )
+    {
+ for( ip in 1996:2016 )
+    {
+    mp <- glm( cbind(X,N-X) ~ Ns(A,knots=a.kn),
+               family = binomial( link=log ),
+                 data = subset( prv, sex==sx & P==ip ) )
+    parr[tp,sx,,paste(ip),] <- ci.pred( mp, newdata=data.frame(A=a.pt) )
+    }
+    mt <- glm( cbind(X,N-X) ~ Ns(A,knots=a.kn) + P,
+               family = binomial( link=log ),
+                 data = subset( prv, sex==sx ) )
+    tarr[tp,sx,] <- ci.exp( mt, subset="P" )
+    }
+    }
> round( range( parr, na.rm=T ), 4 )
```

```
[1] 0.0000 0.1873
> round( ftable( (tarr-1)*100 ), 2 )
          Est   lo   hi
typ sex
T1  M    0.58 0.52 0.64
    F    0.67 0.61 0.74
T2  M    5.21 5.18 5.24
    F    5.03 5.00 5.05
DM  M    4.43 4.41 4.46
    F    4.38 4.35 4.40
```

Thus we see that the overall trend in T1D prevalence is very modest, (about 0.63%/year) as compared to the change in T2 prevalence (about 5.1%/year).

We will need to be able to show only every 5th date of prevalence so we need:

```
> wh <- paste(seq(1996,2016,5))
```

The prevalences for each of the 21 dates are shown in figure 2.1:

```
> par( mfrow=c(2,2), mar=c(3,3,0,0), oma=c(0,0,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> for( tp in dimnames(parr)[["typ"]][1:2] )
+ for( sx in dimnames(parr)[["sex"]] )
+    {
+    matplot( a.pt, parr[tp,sx,,,1]*100,
+             type="l", lwd=c(1,2), lty=1,
+             col=if(sx=="M") "blue" else "red",
+             xlab="", ylab="", xlim=c(5,95), ylim=c(0,if(tp=="T1") 1 else 19),
+             yaxs="i", xaxt="n", yaxt="n" )
+    axis( side=1, at=1:9*10 )
+    axis( side=1, at=1:19*5, tcl=-0.3, labels=NA )
+    if( sx=="M" )
+      {
+      axis(side=2)
+      text( 5, par("usr")[4], tp, font=2, adj=c(0,1) )
+      }
+    }
> mtext( "Age", side=1, line=2, outer=TRUE )
```

```
> p22 <- function(){
+ par( mfrow=c(2,2), mar=c(3,0,0,0), oma=c(0,3,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ for( tp in dimnames(parr)[["typ"]][1:2] )
+ for( sx in dimnames(parr)[["sex"]] )
+    {
+    matplot( a.pt, parr[tp,sx,,wh,1]*100,
+             type="l", lwd=2, lty=1,
+             col=if(sx=="M") "blue" else "red",
+             xlab="", ylab="", xlim=c(5,95), ylim=c(0,if(tp=="T1") 1 else 19),
+             yaxs="i", xaxt="n", yaxt="n" )
+    for( i in wh )
+       polygon( c(a.pt,rev(a.pt)), c(parr[tp,sx,,i,2],
+                                   rev(parr[tp,sx,,i,3]))*100,
+                 col=if(sx=="M") "#0000FF33" else "#FF000033", border="transparent" )
+    axis( side=1, at=1:9*10 )
+    axis( side=1, at=1:19*5, tcl=-0.3, labels=NA )
+    if( tp=="T1" ) {
+    text( rep(60,5), txpos <- seq(0.4,0.6,,5)-0.2*(sx=="F"), wh, adj=0 )
```

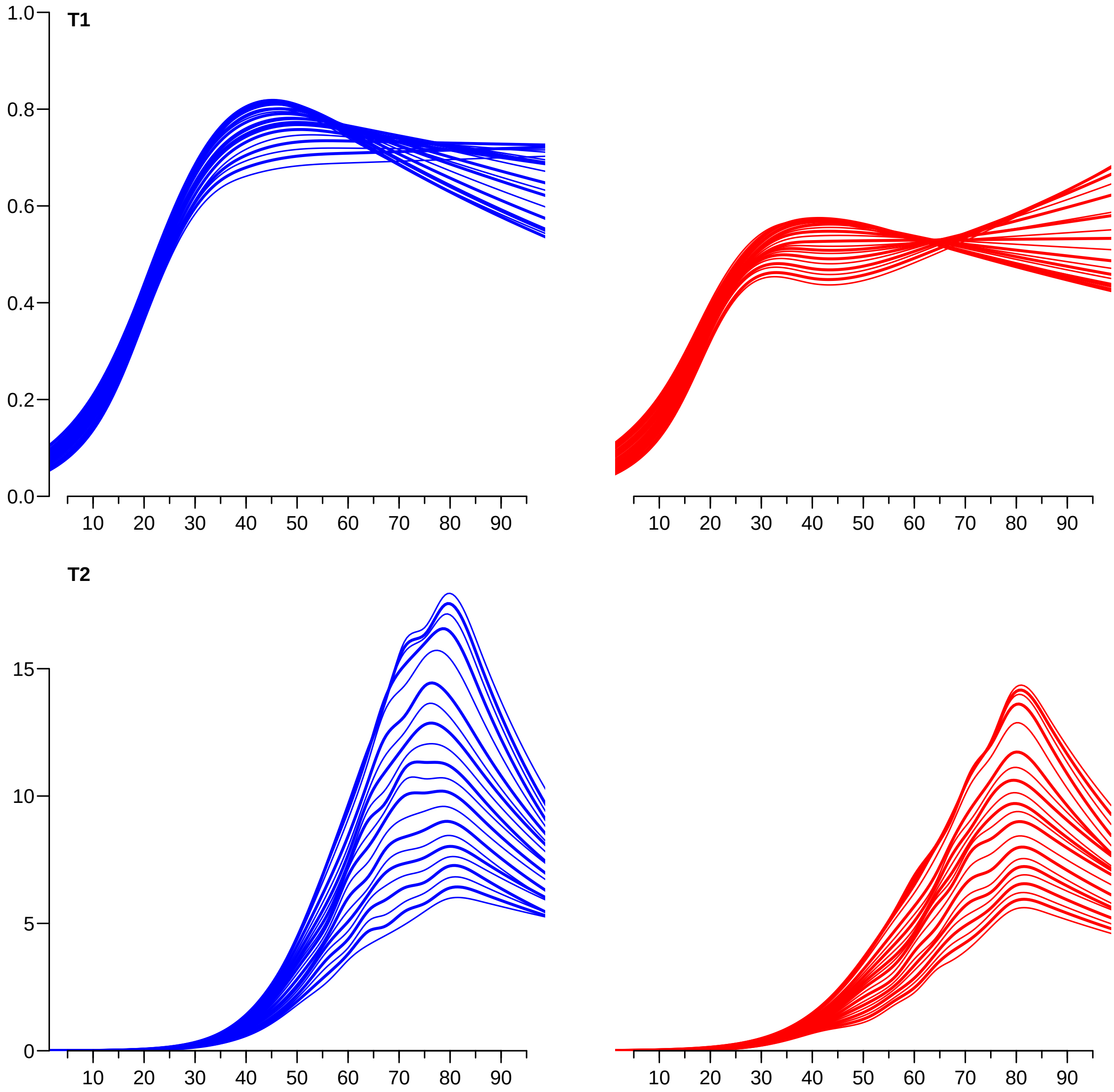

Figure 2.1: *Age-specific prevalences for each of the dates 1096-01-01, 1997-01-01, ... 2016-01-01 for T1 and T2. Blue curves are men, red women. Thin curves are even years, thick odd years.* `./prev-3tp2sx`

```
+    segments( rep(50,5), parr[tp,sx,"50",wh,1]*100,
+              rep(60,5), txpos ) } else {
+    text( rep(90,5), parr[tp,sx,"90",wh,1]*100, wh, cex=0.9, adj=c(1,1) ) }
+    if( sx=="M" )
+      {
+      axis( side=2)
+      axis( side=2, at=if( tp=="T1" ) 1:9/10 else 1:19, labels=NA, tcl=-0.3 )
+      }
+    if( sx=="M" ) mtext( paste( tp, "diabetes prevalence (%)"), side=2, line=2, las=0 )
+    }
+ mtext( "Age", side=1, line=2, outer=TRUE )
```

```
+ mtext( c("Men","Women"), at=c(1,3)/4, side=3, line=0, outer=TRUE )
+ mtext( c("Age","Age"), at=c(1,3)/4, side=1, line=-1, outer=TRUE )
+ }
> p22()
```

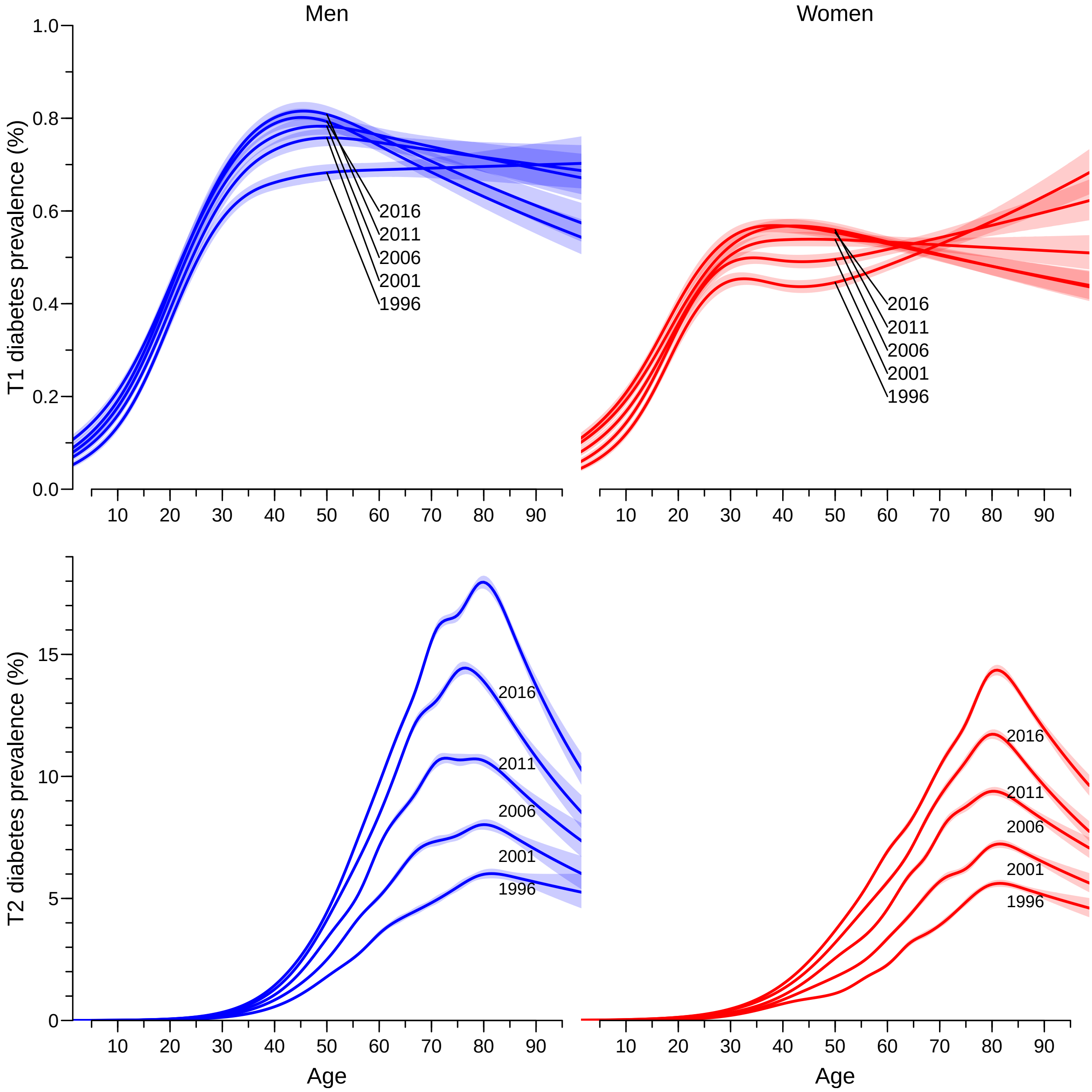

Figure 2.2: *Age-specific prevalences for each of the dates 1996-01-01, 2001-01-01, . . . 1996-01-01 for T1 and T2 diabetes. Blue curves are men, red women, shaded areas are 95% confidence intervals.* `./prev-2tp2sx`

```
> win.metafile( "2tp2sx.emf", width=8, height=8 )
> p22()
> dev.off()
null device
          1
```

We can also present the relative fraction of all diabetes that is T1:

```
> par( mfrow=c(1,2), mar=c(3,3,0,0), oma=c(0,0,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> for( sx in dimnames(parr)[["sex"]] )
+    {
+    matplot( a.pt, parr["T1",sx,,,1]/(parr["T1",sx,,,1]+parr["T2",sx,,,1])*100,
+             type="l", lwd=c(1,2), lty=1,
+             col=if(sx=="M") "blue" else "red",
+             xlab="", ylab="", xlim=c(5,95), ylim=c(0,100),
+             yaxs="i", xaxt="n", yaxt="n" )
+    axis( side=1, at=1:9*10 )
+    axis( side=1, at=1:19*5, tcl=-0.3, labels=NA )
+    if( sx=="M" ) axis(side=2)
+    abline( h=50, v=35, col=gray(0.7) )
+    }
> mtext( "Age", side=1, line=2, outer=TRUE )
```

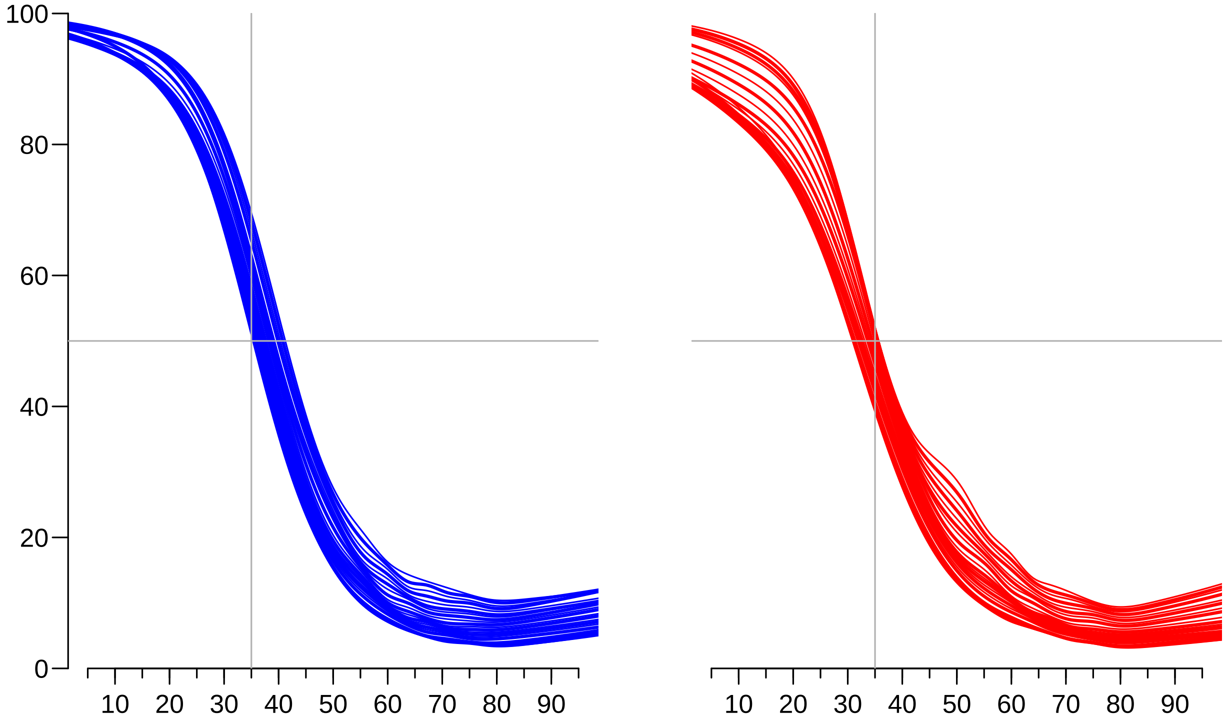

Figure 2.3: *Fraction of T1D among all diabetes patients for dates 1996-01-01, 1997-01-01 etc.* `./prev-p-ratio`

```
> par( mfrow=c(1,2), mar=c(3,3,0,0), oma=c(0,0,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> for( sx in dimnames(parr)[["sex"]] )
+    {
+    matplot( a.pt, crv <- parr["T1",sx,,wh,1]/(parr["T1",sx,,wh,1]+parr["T2",sx,,wh,1])*10
+             type="l", lwd=c(1,2), lty=1,
+             col=if(sx=="M") "blue" else "red",
+             xlab="", ylab="", xlim=c(5,95), ylim=c(0,100),
+             yaxs="i", xaxt="n", yaxt="n" )
+    axis( side=1, at=1:9*10 )
+    axis( side=1, at=1:19*5, tcl=-0.3, labels=NA )
```

```
+     if( sx=="M" ) axis(side=2)
+     abline( h=50, v=35, col=gray(0.7) )
+     text( rep(50,5), crv["50",], wh, cex=0.5, adj=1 )
+     }
> mtext( "Age", side=1, line=2, outer=TRUE )
```

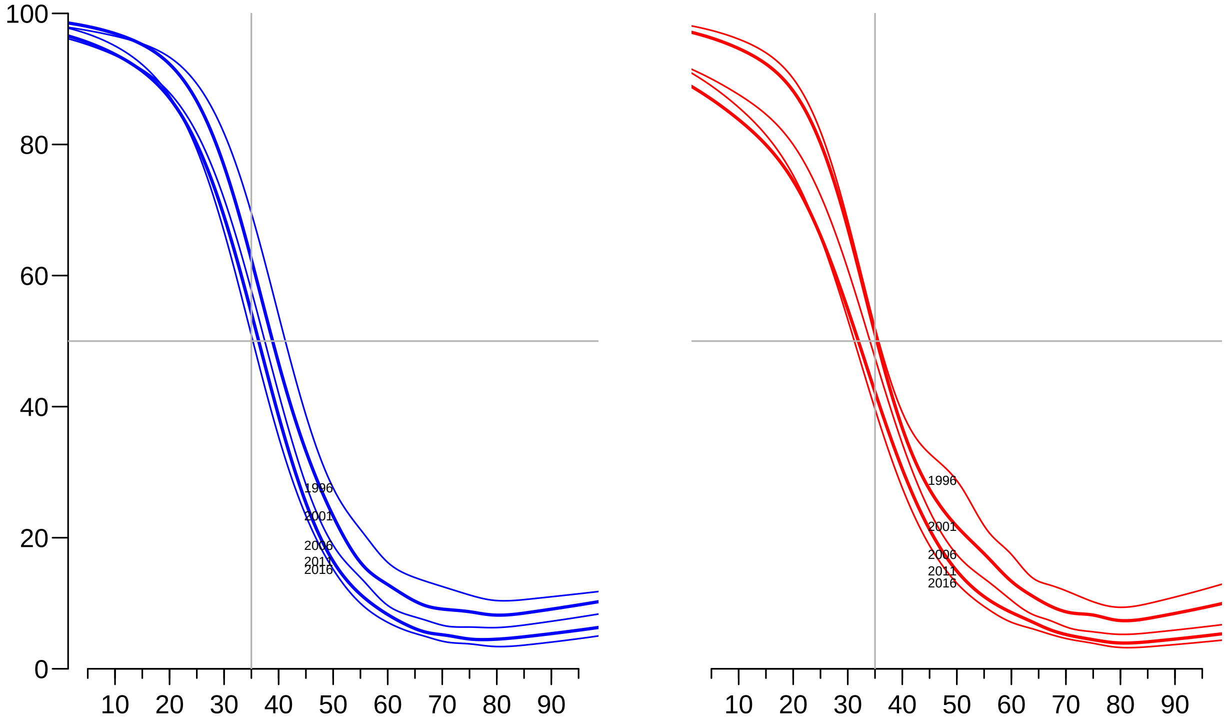

Figure 2.4: *Fraction of T1D among all diabetes patients for dates 1996-01-01,2001-01-01 etc. In both panels, at age 50 the upper curve is for 1996, and the lower for 2016, reflecting the steeper increase in T2D.* `./prev-p5-ratio`

. . . now input from `rates.tex`

Chapter 3

Rate data

We read the register-derived datasets — the `rtL` is the dataset with only `L`ate diagnoses of DM, that is after 1995-12-31, so the part of the follow-up where we are reasonably sure about the data of diagnosis even in the absence of the DADD information.

```
> rt  <- read_sas("../data/FUtot.sas7bdat")
> rtL <- read_sas("../data/FUtotL.sas7bdat")
```

The datasets are primarily classified by `state`, the state *from* which events (`T1`, `T2` or `D`) may occur.

```
> fCtable( xtabs( cbind( T1, T2, D, Y ) ~ state + dur, data=subset(rt,dur<5) ),
+          row.vars=2, w=7 )
    state    noDM                              T1                              T2
               T1      T2       D       Y      T1      T2       D       Y      T1      T2
dur
0          19,803 228,670 998,533 110,801       .       .   1,803      98       .       .
0.2             .       .       .       .       .       .     303       7       .       .
0.5             .       .       .       .       .       .     356      12       .       .
1               .       .       .       .       .       .     577      24       .       .
2               .       .       .       .       .       .     588      24       .       .
3               .       .       .       .       .       .     615      23       .       .
4               .       .       .       .       .       .     667      23       .       .
> with( rt, table(P-A-C) )

     0      1
215459 205097
```

So, the age-period-cohort classification is in Lexis triangles so we must devise the proper age- and period-midpoints (and consequently cohort-midpoints). Moreover, we also recode the `dur` (duration of diabetes) to the midpoints of the duration intervals:

```
> rt <- transform( as.data.frame(subset( rt, P>1995.5 & Y>1e-6 )),
+                  Ax = A + (1+ P-A-C )/3,
+                  Px = P + (2-(P-A-C))/3,
+                 sex = factor( sex, labels=c("M","F") ),
+               state = factor( state ) )
> rt <- transform( rt, Cx = Px-Ax,
+                     dur = ifelse( state=="noDM", 0,
+                                   dur + 0.5 - 0.25*(dur<0.9)
+                                             - 0.10*(dur<0.4)
+                                             - 0.05*(dur<0.1) ) )
> with( subset(rt, Ax<5 & Px<1999), print( table(round(Ax,2),round(Px,2)), z="." ) )
```

```
       1996.33 1996.67 1997.33 1997.67 1998.33 1998.67
  0.33       .       8       .       5       .       5
  0.67      11       .       9       .       7       .
  1.33       .      13       .      12       .      12
  1.67      11       .      14       .      15       .
  2.33       .      13       .      16       .      14
  2.67      15       .      12       .      15       .
  3.33       .      17       .      15       .      18
  3.67      19       .      17       .      16       .
  4.33       .      22       .      21       .      18
  4.67      21       .      21       .      21       .
> fCtable( xtabs( cbind( T1, T2, D, Y ) ~ state + dur, data=subset(rt,dur<5) ),
+          row.vars=2, w=7 )
     state    noDM                                T1                              T2
                T1      T2       D       Y      T1      T2       D       Y      T1      T2
dur
0           17,329 189,920 941,489 105,629       .       .       .       .       .       .
0.1              .       .       .       .       .       .   1,781      96       .       .
0.35             .       .       .       .       .       .     279       6       .       .
0.75             .       .       .       .       .       .     314      11       .       .
1.5              .       .       .       .       .       .     511      22       .       .
2.5              .       .       .       .       .       .     523      22       .       .
3.5              .       .       .       .       .       .     565      22       .       .
4.5              .       .       .       .       .       .     607      22       .       .
```

. . . and the same with the Late data frame:

```
> rtL <- transform( as.data.frame(subset( rtL, P>1995.5 & Y>1e-6 )),
+                    Ax = A + (1+ P-A-C )/3,
+                    Px = P + (2-(P-A-C))/3,
+                   sex = factor( sex, labels=c("M","F") ),
+                 state = factor( state ) )
> rtL <- transform( rtL, Cx = Px-Ax,
+                        dur = ifelse( state=="noDM", 0,
+                                      dur + 0.5 - 0.25*(dur<0.9)
+                                                - 0.10*(dur<0.4)
+                                                - 0.05*(dur<0.1) ) )
> with( subset(rtL, Ax<5 & Px<1999), print( table(round(Ax,2),round(Px,2)), z="." ) )
       1996.33 1996.67 1997.33 1997.67 1998.33 1998.67
  0.33       .       8       .       5       .       5
  0.67       8       .       9       .       7       .
  1.33       .       9       .      12       .      12
  1.67       8       .      13       .      15       .
  2.33       .      10       .      13       .      14
  2.67       7       .      10       .      14       .
  3.33       .      10       .      12       .      15
  3.67       8       .      10       .      14       .
  4.33       .      10       .      14       .      15
  4.67      11       .      10       .      14       .
> fCtable( xtabs( cbind( T1, T2, D, Y ) ~ state + dur, data=subset(rtL,dur<5) ),
+                row.vars=2, w=7 )
     state    noDM                                T1                              T2
                T1      T2       D       Y      T1      T2       D       Y      T1      T2
dur
0           17,329 189,920 941,489 105,629       .       .       .       .       .       .
0.1              .       .       .       .       .       .   1,780      96       .       .
```

```
0.35             .       .       .       .       .       .     273       6       .       .
0.75             .       .       .       .       .       .     303      10       .       .
1.5              .       .       .       .       .       .     421      20       .       .
2.5              .       .       .       .       .       .     355      18       .       .
3.5              .       .       .       .       .       .     351      17       .       .
4.5              .       .       .       .       .       .     331      16       .       .
```

```
> summary( rtL )
```

```
 sex        state             A               P              C             dur        
 M:93469   noDM: 8000   Min.   : 0.00   Min.   :1996   Min.   :1896   Min.   : 0.000  
 F:95374   T1   :90893   1st Qu.:32.00   1st Qu.:2004   1st Qu.:1934   1st Qu.: 0.750  
           T2   :89950   Median :52.00   Median :2009   Median :1955   Median : 3.500  
                        Mean   :52.11   Mean   :2008   Mean   :1955   Mean   : 5.177  
                        3rd Qu.:73.00   3rd Qu.:2012   3rd Qu.:1976   3rd Qu.: 8.500  
                        Max.   :99.00   Max.   :2015   Max.   :2015   Max.   :19.500  
       D                T1                 T2                 Y            
 Min.   :  0.00   Min.   : 0.00000   Min.   :  0.000   Min.   : 0.000001  
 1st Qu.:  0.00   1st Qu.: 0.00000   1st Qu.:  0.000   1st Qu.: 0.001192  
 Median :  0.00   Median : 0.00000   Median :  0.000   Median : 0.003530  
 Mean   :  5.53   Mean   : 0.09176   Mean   :  1.006   Mean   : 0.572999  
 3rd Qu.:  1.00   3rd Qu.: 0.00000   3rd Qu.:  0.000   3rd Qu.: 0.019543  
 Max.   :587.00   Max.   :17.00000   Max.   :279.000   Max.   :23.341741  
       Ax                Px             Cx      
 Min.   : 0.3333   Min.   :1996   Min.   :1897  
 1st Qu.:32.3333   1st Qu.:2004   1st Qu.:1935  
 Median :52.6667   Median :2009   Median :1955  
 Mean   :52.6048   Mean   :2008   Mean   :1956  
 3rd Qu.:73.6667   3rd Qu.:2013   3rd Qu.:1976  
 Max.   :99.6667   Max.   :2016   Max.   :2015  
```

We then save the groomed dataframes:

```
> save( rt, rtL, file="../data/rt.Rda" )
```

Chapter 4

DM incidence

In this chapter we use the constructed follow-up dataset for the entire population:

```
> options( width=95 )
> library( Epi )
> library( splines )
> clear()
> load( file="../data/inits.Rda" )
> load( file="../data/rt.Rda" )
```

Modeling incidence rates of diabetes will involve only person in the "noDM" state, which is the same from the two datsets.

We restict data to the `noDM` state, and show the number of cases and person-years in slightly different guises:

```
> str( st <- subset( rt , state=="noDM" ) )
'data.frame':        8000 obs. of  13 variables:
 $ sex  : Factor w/ 2 levels "M","F": 1 1 1 1 1 1 1 1 1 1 ...
 $ state: Factor w/ 3 levels "noDM","T1","T2": 1 1 1 1 1 1 1 1 1 1 ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ P    : num  1996 1996 1997 1997 1998 ...
 $ C    : num  1995 1996 1996 1997 1997 ...
 $ dur  : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  28 0 18 1 18 2 11 0 21 0 ...
 $ T1   : num  1 0 1 1 1 2 2 2 3 1 ...
 $ T2   : num  0 0 0 0 1 0 0 0 0 0 ...
 $ Y    : num  17.9 17.6 17.3 17.7 17.1 ...
 $ Ax   : num  0.667 0.333 0.667 0.333 0.667 ...
 $ Px   : num  1996 1997 1997 1998 1998 ...
 $ Cx   : num  1996 1996 1997 1997 1998 ...
> fCtable( addmargins( xtabs( cbind(T1,T2) ~ P + sex, data=st  ),
+                      margin=2:3 ),
+          col.vars=3:2, w=6 )
            T1                       T2                       Sum
     sex     M      F    Sum      M      F    Sum      M      F    Sum
P
1996       551    389    940  3,315  2,463  5,778  3,866  2,852  6,718
1997       488    393    881  3,375  2,475  5,850  3,863  2,868  6,731
1998       513    346    859  3,640  2,604  6,244  4,153  2,950  7,103
1999       491    350    841  4,042  2,989  7,031  4,533  3,339  7,872
2000       516    371    887  3,976  2,889  6,865  4,492  3,260  7,752
2001       484    328    812  3,966  2,912  6,878  4,450  3,240  7,690
```

```
2002        463   334   797  4,623  3,592  8,215  5,086  3,926  9,012
2003        456   328   784  5,183  4,022  9,205  5,639  4,350  9,989
2004        464   329   793  5,204  4,007  9,211  5,668  4,336 10,004
2005        560   419   979  5,534  4,185  9,719  6,094  4,604 10,698
2006        562   395   957  5,562  3,990  9,552  6,124  4,385 10,509
2007        549   381   930  5,907  4,474 10,381  6,456  4,855 11,311
2008        562   378   940  6,411  4,758 11,169  6,973  5,136 12,109
2009        565   380   945  6,893  4,885 11,778  7,458  5,265 12,723
2010        572   380   952  7,301  5,331 12,632  7,873  5,711 13,584
2011        492   383   875  8,796  7,116 15,912  9,288  7,499 16,787
2012        479   309   788  7,456  5,722 13,178  7,935  6,031 13,966
2013        463   342   805  5,689  4,308  9,997  6,152  4,650 10,802
2014        402   341   743  5,539  4,022  9,561  5,941  4,363 10,304
2015        471   350   821  6,165  4,599 10,764  6,636  4,949 11,585
```

```
> fCtable( addmargins( xtabs( cbind(T1,T2,Y) ~ P + sex, data=st  ),
+                      margin=2 ),
+          col.vars=3:2, w=6 )
             T1                  T2                   Y
     sex      M     F   Sum       M      F    Sum      M      F    Sum
P
1996        551   389   940  3,315  2,463  5,778  2,550  2,617  5,167
1997        488   393   881  3,375  2,475  5,850  2,558  2,624  5,182
1998        513   346   859  3,640  2,604  6,244  2,563  2,630  5,193
1999        491   350   841  4,042  2,989  7,031  2,568  2,634  5,202
2000        516   371   887  3,976  2,889  6,865  2,573  2,640  5,212
2001        484   328   812  3,966  2,912  6,878  2,579  2,646  5,225
2002        463   334   797  4,623  3,592  8,215  2,585  2,650  5,235
2003        456   328   784  5,183  4,022  9,205  2,588  2,652  5,240
2004        464   329   793  5,204  4,007  9,211  2,589  2,654  5,243
2005        560   419   979  5,534  4,185  9,719  2,592  2,657  5,249
2006        562   395   957  5,562  3,990  9,552  2,598  2,662  5,260
2007        549   381   930  5,907  4,474 10,381  2,607  2,670  5,277
2008        562   378   940  6,411  4,758 11,169  2,620  2,681  5,302
2009        565   380   945  6,893  4,885 11,778  2,632  2,693  5,326
2010        572   380   952  7,301  5,331 12,632  2,641  2,703  5,344
2011        492   383   875  8,796  7,116 15,912  2,648  2,710  5,358
2012        479   309   788  7,456  5,722 13,178  2,654  2,715  5,369
2013        463   342   805  5,689  4,308  9,997  2,665  2,724  5,389
2014        402   341   743  5,539  4,022  9,561  2,683  2,739  5,422
2015        471   350   821  6,165  4,599 10,764  2,692  2,741  5,434
```

Note that the `Y` is in units of 1000 PY.

4.1 Age-Period-Cohort models

We fit APC-models for T1, T2 and all diabetes incidence separately for men and women and plot the estimates together for men and women:

```
> ( A1.kn <- with( st, quantile( rep(    Ax, T1 ), qn(5) ) ) )
      10%       30%       50%       70%       90%
 7.333333 14.666667 26.333333 44.333333 58.666667
```

```
> ( P1.kn <- with( st, quantile( rep( Px    , T1 ), qn(4) ) ) )
   12.5%    37.5%    62.5%    87.5%
1998.667 2003.667 2008.667 2013.333
```

```
> ( C1.kn <- with( st, quantile( rep( Px-Ax, T1 ), qn(5) ) ) )
     10%      30%      50%      70%      90% 
1944.667 1964.333 1978.667 1991.333 2001.333 
> ( A2.kn <- with( st, quantile( rep(    Ax, T2 ), qn(7) ) ) )
7.142857% 21.42857% 35.71429%       50% 64.28571% 78.57143% 92.85714% 
 38.66667  47.66667  52.33333  56.33333  59.33333  62.33333  74.33333 
> ( P2.kn <- with( st, quantile( rep( Px    , T2 ), qn(5) ) ) )
     10%      30%      50%      70%      90% 
1999.333 2004.333 2008.333 2011.333 2014.333 
> ( C2.kn <- with( st, quantile( rep( Px-Ax, T2 ), qn(7) ) ) )
7.142857% 21.42857% 35.71429%       50% 64.28571% 78.57143% 92.85714% 
 1931.667  1942.667  1947.333  1951.333  1955.333  1961.667  1971.667 
```

With these knots (similar for men and women) we can fit 4 different APC-models — for men and women and T1D and T2D. We devise

```
> M1 <- apc.fit( transform( subset(st,sex=="M"), A=Ax, P=Px, D=T1 )[,c("A","P","D","Y")],
+                 parm="APC", ref.p=2010, npar=list(A=A1.kn,P=P1.kn,C=C1.kn), scale=100 )
[1] "ML of APC-model Poisson with log(Y) offset : ( APC ):\n"

Analysis of deviance for Age-Period-Cohort model

                  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)    
Age                    3995     7430.8                          
Age-drift              3994     7425.2  1    5.563   0.01834 *  
Age-Cohort             3991     7161.0  3  264.210 < 2.2e-16 ***
Age-Period-Cohort      3989     7076.4  2   84.683 < 2.2e-16 ***
Age-Period             3992     7372.0 -3 -295.634 < 2.2e-16 ***
Age-drift              3994     7425.2 -2  -53.258 2.723e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> F1 <- apc.fit( transform( subset(st,sex=="F"), A=Ax, P=Px, D=T1 )[,c("A","P","D","Y")],
+                 parm="APC", ref.p=2010, npar=list(A=A1.kn,P=P1.kn,C=C1.kn), scale=100 )
[1] "ML of APC-model Poisson with log(Y) offset : ( APC ):\n"

Analysis of deviance for Age-Period-Cohort model

                  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)    
Age                    3995     6170.5                          
Age-drift              3994     6167.1  1    3.413  0.064686 .  
Age-Cohort             3991     6002.3  3  164.785 < 2.2e-16 ***
Age-Period-Cohort      3989     5978.4  2   23.993 6.165e-06 ***
Age-Period             3992     6154.1 -3 -175.752 < 2.2e-16 ***
Age-drift              3994     6167.1 -2  -13.027  0.001483 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> M2 <- apc.fit( transform( subset(st,sex=="M"), A=Ax, P=Px, D=T2 )[,c("A","P","D","Y")],
+                 parm="APC", ref.p=2010, npar=list(A=A2.kn,P=P2.kn,C=C2.kn) )
[1] "ML of APC-model Poisson with log(Y) offset : ( APC ):\n"

Analysis of deviance for Age-Period-Cohort model

                  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)    
```

```
Age                    3993      25254
Age-drift              3992      21428  1   3825.6 < 2.2e-16 ***
Age-Cohort             3987      19968  5   1460.3 < 2.2e-16 ***
Age-Period-Cohort      3984      17698  3   2269.6 < 2.2e-16 ***
Age-Period             3989      19655 -5  -1956.5 < 2.2e-16 ***
Age-drift              3992      21428 -3  -1773.3 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> F2 <- apc.fit( transform( subset(st,sex=="F"), A=Ax, P=Px, D=T2 )[,c("A","P","D","Y")],
+                parm="APC", ref.p=2010, npar=list(A=A2.kn,P=P2.kn,C=C2.kn) )
[1] "ML of APC-model Poisson with log(Y) offset : ( APC ):\n"

Analysis of deviance for Age-Period-Cohort model

                  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
Age                    3993      26818
Age-drift              3992      23573  1   3244.8 < 2.2e-16 ***
Age-Cohort             3987      22015  5   1558.5 < 2.2e-16 ***
Age-Period-Cohort      3984      20003  3   2011.4 < 2.2e-16 ***
Age-Period             3989      22079 -5  -2075.1 < 2.2e-16 ***
Age-drift              3992      23573 -3  -1494.8 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We stick to the period-major parametrization as in figure **??**, we are essentially referring to cross-sectional rates, and they seem to have a peak around age 60 — which is a bit fishy…

```
> par( mfrow=c(2,1), mar=c(0,4,0,4), oma=c(3,0,1,0), mgp=c(3,1,0)/1.6, las=1 )
> apc.frame( a.lab=seq(0,90,20), cp.lab=seq(1900,2020,30), r.lab=c(c(1,2,5),c(1,2,5)*10),
+            a.tic=seq(0,90,10), cp.tic=seq(1900,2020,10), r.tic=c(1:10,1:5*10),
+            gap=17, r.txt="T1 diabetes incidence rate per 100,000 PY", side=c(2,4) )
> apc.lines( M1, col="blue" )
> apc.lines( F1, col="red" )
> apc.frame( a.lab=seq(0,90,20), cp.lab=seq(1900,2020,30), r.lab=c(c(1,2,5)/100,c(1,2,5)/10
+            a.tic=seq(0,90,10), cp.tic=seq(1900,2020,10), r.tic=c(1:10/100,1:10/10,1:15),
+            gap=17, r.txt="T2 diabetes incidence rate per 1000 PY", side=c(1,2,4) )
> apc.lines( M2, col="blue" )
> apc.lines( F2, col="red" )
```

4.1.1 Time-trends in rates

We can get a raw overview of the average trend in incidence, although the period effcts in figure **??** are very curved, so a linear average is not worth much:

```
> M1$Drift
                exp(Est.)      2.5%     97.5%
APC (D-weights) 0.9938661 0.9902135 0.9975321
A-d             0.9959378 0.9925748 0.9993122
> dr <- rbind(
+ cbind( M1$Drift, M2$Drift )[1,],
+ cbind( F1$Drift, F2$Drift )[1,] )
> rownames( dr ) <- c("M","F")
> colnames( dr )[c(1,4)] <- c("T1","     T2")
> round( (dr-1)*100, 1 )
    T1 2.5% 97.5%      T2 2.5% 97.5%
M -0.6 -1.0  -0.2     3.9  3.8   4.1
F -1.1 -1.5  -0.6     3.8  3.7   4.0
```

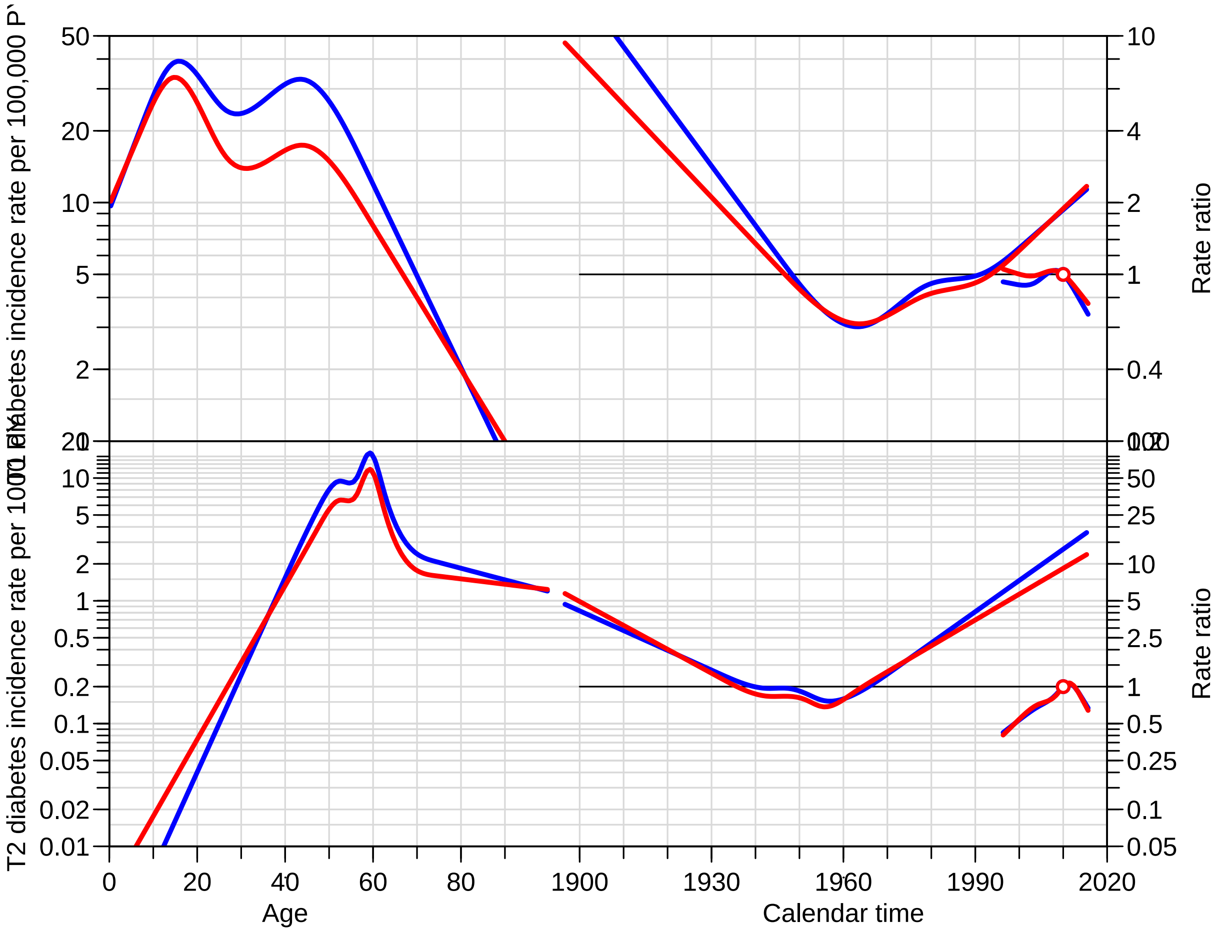

Figure 4.1: *Estimated APC-effects for T1 and T2 diabetes for men (blue) and women (red).* `./inc-apc-pl`

4.1.2 Predicted rates

To that end we use the model-objects returned by the `apc.fit` function to produce predicted rates. So we set up a prediction frame with ages for 15 different cohorts:

```
> prf <- data.frame( A = rep( c(NA,0:98), 8 ),
+                    C = rep( seq(1910,1980,10), each=100 ),
+                    Y = 1 )[-1,]
> prf <- transform( prf, P = C + A )
```

The we can make a fit of the models of relevance and make predictions based on this new frame. [1]

```
> Mapc <- glm( D ~ Ns(   A, kn=apcM$Knots$Age ) +
+                  Ns( P-A, kn=apcM$Knots$Coh ) +
+                  Ns( P  , kn=apcM$Knots$Per ),
+              offset = log( Y ),
+              family = poisson,
+                data = subset( DD, sex=="M") )
> Map  <- glm( D ~ Ns( A, kn=apcM$Knots$Age ) +
```

---

[1] Note that we cannot use the returned model from the `apc` object since this is defined in terms specific matrices and *not* in terms of A, P and C:

```
+                  Ns( P, kn=apcM$Knots$Per ),
+              offset = log( Y ),
+              family = poisson,
+                data = subset( DD, sex=="M") )
> Mac  <- glm( D ~ Ns(   A, kn=apcM$Knots$Age ) +
+                  Ns( P-A, kn=apcM$Knots$Coh ),
+              offset = log( Y ),
+              family = poisson,
+                data = subset( DD, sex=="M") )
> Fapc <- glm( D ~ Ns(   A, kn=apcF$Knots$Age ) +
+                  Ns( P-A, kn=apcF$Knots$Coh ) +
+                  Ns( P  , kn=apcF$Knots$Per ),
+              offset = log( Y ),
+              family = poisson,
+                data = subset( DD, sex=="F") )
> Fap  <- glm( D ~ Ns( A, kn=apcF$Knots$Age ) +
+                  Ns( P, kn=apcF$Knots$Per ),
+              offset = log( Y ),
+              family = poisson,
+                data = subset( DD, sex=="F") )
> Fac  <- glm( D ~ Ns(   A, kn=apcF$Knots$Age ) +
+                  Ns( P-A, kn=apcF$Knots$Coh ),
+              offset = log( Y ),
+              family = poisson,
+                data = subset( DD, sex=="F") )
> summary( fitted( apcM$Model ) - fitted( Mapc ) )
> summary( fitted( apcF$Model ) - fitted( Fapc ) )
```

From the last summary we see that the models are the same as those fitted by `apc.fit`, an moreover we can use this latter to make predictions, regardless of the overparametrization (we will get a warning, though). Recall that the `Y` was scaled to be person-millenia, so we get fitted values as rates per 1000 (namely the expected numbers based on the model for a data point where `Y` is equal to 1, as specified in `prf`):

```
> prr <- subset( prf, (P<2016 & P>1995) | is.na(P) )
> Mfit.apc <- predict( Mapc, newdata=prr )
> Mfit.ap  <- predict( Map , newdata=prr )
> Mfit.ac  <- predict( Mac , newdata=prr )
> Ffit.apc <- predict( Fapc, newdata=prr )
> Ffit.ap  <- predict( Fap , newdata=prr )
> Ffit.ac  <- predict( Fac , newdata=prr )
```

For comparison we overlay empirical rates, which we compute for the cohorts 1910 (born 1905–15), . . . , 1980 (born 1975–85) calculated in C-sets (✎); the `gc` and `gp` are the midpoints of the cohort and period in the C-sets:

```
> DD.x <- transform( DD,
+                    gc = floor(((P-A)-1905)/10)*10+1910,
+                    gp = floor(P)+0.5 )
> ee <- data.frame( xtabs( cbind(D,Y) ~ sex + gp + gc,
+                   data = subset( DD.x, gc>1905 & gc<1985 ) ) )
> ee <- reshape( ee, timevar = "Var4",
+                      idvar = c("sex","gp","gc"),
+                        dir = "wide" )
> names( ee )[4:5] <- c("D","Y")
> ee <- transform( ee, gp = as.numeric(as.character(gp)),
+                      gc = as.numeric(as.character(gc)) )
> str( ee )
```

We then overlay the empirical over the fitted rates from the three different models, the age-period, the age-cohort and the apc-model:

```
> par( mfrow=c(2,1), mar=c(0,0,0,0), oma=c(3,3,1,1), mgp=c(3,1,0)/1.6 )
> matplot( prr$A, exp(Mfit.apc), type="l", lty=1,
+          log="y", ylim=c(0.2,25), lwd=3, xaxt="n", xlab="", ylab="" )
> matlines( prr$A, exp(Mfit.ap), type="l", lty="11", lwd=2 )
> matlines( prr$A, exp(Mfit.ac), type="l", lty=1, lwd=2 )
> with( subset(ee,sex=="M"),
+       points( gp-gc, D/Y, pch=16, col=rainbow(8)[factor(gc)], cex=0.8 ) )
> text( 20, 14, "Men", col="blue" )
> matplot( prr$A, exp(Ffit.apc), type="l", lty=1,
+          log="y", ylim=c(0.2,25), lwd=3, xlab="", ylab="" )
> matlines( prr$A, exp(Ffit.ap), type="l", lty="11", lwd=2 )
> matlines( prr$A, exp(Ffit.ac), type="l", lty=1, lwd=2 )
> with( subset(ee,sex=="F"),
+       points( gp-gc, D/Y, pch=16, col=rainbow(8)[factor(gc)], cex=0.8 ) )
> text( 20, 14, "Women", col="red" )
> mtext( "DM incidence rate per 1000 PY", side=2, outer=TRUE, line=2, las=0 )
> mtext( "Age (years)", side=1, outer=TRUE, line=2 )
```

4.1.3 Prediction by age across periods

Alternatively we could show the rates in different ages as a function of calendar time; we shall do this for ages 20, 30,. . . ,90, note that we use the natural spline property of linearity to boldly predict rates beyond 2016-01-01.

```
> prf <- data.frame( A = rep( 2:8*10, each=31 ),
+                    P = rep( 1995:2025, 7 ),
+                    Y = 1 )[outer(c(NA,1:31),0:6*31,"+")[-1],]
```

Then we can show the predicted rates:

```
> plp <-
+ function( mod, clr, pr, ya="l" )
+     {
+ matplot( prf$P, ci.pred( mod, prf ),
+          type="l", lty=c("solid","22","66","11","33")[c(1,1,1,1,3,3,4,5,5)],
+          lwd=c(3,1,1), col=clr,
+          log="y", xaxt="n", xlab="", ylab="", ylim=c(0.2,20), yaxt=ya )
+ abline( v = 2016, col="gray" )
+     }
> par( mfcol=c(3,2), mar=c(0,0,0,0), oma=c(3,3,1,1), mgp=c(3,1,0)/1.6 )
> plp( Mapc, "blue", prf ) ; text( 1995, 20, "APC", adj=c(0,1) )
> plp( Map , "blue", prf ) ; text( 1995, 20, "AP" , adj=c(0,1) )
> plp( Mac , "blue", prf ) ; text( 1995, 20, "AC" , adj=c(0,1) )
> axis( side=1 )
> plp( Fapc, "red", prf, "n" )
> plp( Fap , "red", prf, "n" )
> plp( Fac , "red", prf, "n" )
> axis( side=1 )
> mtext( "DM incidence rate per 1000 PY", side=2, outer=TRUE, line=2, las=0 )
> mtext( "Date of FU", side=1, outer=TRUE, line=2 )
```

4.1.4 Saving the fitted models

We then save these fitted APC-models with different parametrizations:

```
> save( DD, file="../data/incdata.Rda" )
> save( acpM, apcM, acpF, apcF,
+       Mapc, Mac, Map, Fapc, Fac, Fap, file="../data/inc.Rda" )
```

. . . now input from `mort.tex`

Chapter 5

Mortality

First we load the rate datasets:

```
> library( Epi )
> clear()
> load( file="../data/rt.Rda" )
```

We now have the rate dataset(s) that we will use for modeling of mortality, by age, duration of diabetes, age at diagnosis of diabetes (`dur-Ax`) and calendar time, separately for the two sexes, and for T1D, T2D and for all diabetes together.

5.1 Type 1 diabetes patients

The mortality is modeled using only the rows of the dataframe where the person-years is positive (that is where follow-up is present:

```
> rt1 <- subset( rt, state=="T1" )
> dim( rt1 )
[1] 193282     13
```

In this dataset we now model mortality among men and women separately using age, duration and calendar time as covariates. For convenience we use the same set of knots for the splines

```
> ( a.kn <-     with( rt1, quantile( rep(Ax    ,D), (1:8-0.5)/8 ) )  )
   6.25%   18.75%   31.25%   43.75%   56.25%   68.75%   81.25%   93.75% 
44.33333 56.33333 63.66667 69.33333 74.33333 78.66667 83.33333 89.33333 
> ( d.kn <- c(0,with( rt1, quantile( rep(   dur,D), (1:3-0.0)/4 ) )) )
      25%  50%  75% 
 0.0  3.5  9.5 15.5 
> ( e.kn <-     with( rt1, quantile( rep(Ax-dur,D), (1:5-0.5)/5 ) )  )
     10%      30%      50%      70%      90% 
35.16667 51.83333 61.83333 70.16667 79.23333 
> ( p.kn <-     with( rt1, quantile( rep(   Px ,D), (1:5-0.5)/5 ) )  )
     10%      30%      50%      70%      90% 
1998.333 2001.667 2005.667 2009.333 2013.333 
> ( c.kn <-     with( rt1, quantile( rep(Px-Ax ,D), (1:9-0.5)/9 ) )  )
```

```
5.555556% 16.66667% 27.77778% 38.88889%       50% 61.11111% 72.22222% 83.33333% 94.44444%
 1913.667  1920.667  1924.667  1929.333  1934.333  1939.333  1944.667  1951.667  1964.333
```

With these knot-vectors in place we set up models for men and women:

```
> mt0 <- glm( D ~ Ns( Ax    , knots=a.kn ) +
+                 Ns( Px    , knots=p.kn ),
+             offset = log(Y),
+             family = poisson,
+               data = subset( rt1, sex=="M" ) )
> ft0 <- update( mt0, data = subset( rt1, sex=="F" ) )
> mt1 <- glm( D ~ Ns( Ax    , knots=a.kn ) +
+                 Ns(    dur, knots=d.kn ) +
+                 Ns( Ax-dur, knots=e.kn ) +
+                 Ns( Px    , knots=p.kn ),
+             offset = log(Y),
+             family = poisson,
+               data = subset( rt1, sex=="M" ) )
> ft1 <- update( mt1, data = subset( rt1, sex=="F" ) )
> ml1 <- update( mt1, . ~ . - Ns( Px, knots=p.kn ) + Px )
> fl1 <- update( ft1, . ~ . - Ns( Px, knots=p.kn ) + Px )
> anova( mt1, ml1, test="Chisq" )
Analysis of Deviance Table

Model 1: D ~ Ns(Ax, knots = a.kn) + Ns(dur, knots = d.kn) + Ns(Ax - dur,
    knots = e.kn) + Ns(Px, knots = p.kn)
Model 2: D ~ Ns(Ax, knots = a.kn) + Ns(dur, knots = d.kn) + Ns(Ax - dur,
    knots = e.kn) + Px
  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1     95129      36109
2     95132      36156 -3  -47.151 3.229e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> anova( ft1, fl1, test="Chisq" )
```

```
Analysis of Deviance Table

Model 1: D ~ Ns(Ax, knots = a.kn) + Ns(dur, knots = d.kn) + Ns(Ax - dur,
    knots = e.kn) + Ns(Px, knots = p.kn)
Model 2: D ~ Ns(Ax, knots = a.kn) + Ns(dur, knots = d.kn) + Ns(Ax - dur,
    knots = e.kn) + Px
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1     98117      28089
2     98120      28105 -3  -15.306 0.001573 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> round( (ci.exp( ml1, subset="Px" ) - 1)*100, 1 )
```

```
   exp(Est.) 2.5% 97.5%
Px      -1.5 -1.9  -1.1
```

```
> round( (ci.exp( fl1, subset="Px" ) - 1)*100, 1 )
```

```
   exp(Est.) 2.5% 97.5%
Px      -1.7 -2.1  -1.2
```

```
> mr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(predict(mt1,type="response")),
+             family = poisson,
+               data = subset(rt1,sex=="M") )
```

```
> fr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(predict(ft1,type="response")),
+             family = poisson,
+               data = subset(rt1,sex=="F") )
```

We then set up data frames for predicting the mortality rates for men and women at 1.1.2015, for different combinations of age and duration:

```
> e.pr <- seq(10,75,5)
> d.pr <- c( NA, seq(0,20,0.1) )
> nd0 <- data.frame( Ax = seq( 5,90,0.2), Px=2015, Y=1 )
> nd <- data.frame( expand.grid( d.pr, e.pr ) )
> names( nd ) <- c("dur","Ae")
> nd <- transform( nd, Ax = Ae+dur,
+                      Px = 2015,
+                       Y = 1 )
> head( nd )
  dur Ae   Ax   Px Y
1  NA 10   NA 2015 1
2 0.0 10 10.0 2015 1
3 0.1 10 10.1 2015 1
4 0.2 10 10.2 2015 1
5 0.3 10 10.3 2015 1
6 0.4 10 10.4 2015 1
```

The point in having an `NA` in the prediction data frame is that we can plot different lines from one vector — `NA`s in a vector produces a break between the points on either side of the `NA`:

```
> t0pr <- cbind( ci.pred( mt0, nd0 ), ci.pred( ft0, nd0 ) )
> t0pr <- cbind( t0pr, ci.ratio(t0pr[,1:3],t0pr[,1:3+3]) )
> t1pr <- cbind( ci.pred( mt1, nd ), ci.pred( ft1, nd ) )
> t1pr <- cbind( t1pr, ci.ratio(t1pr[,1:3],t1pr[,1:3+3]) )
> par( mar=c(3,3,1,3), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> plot( NA,
+       log="y", xlab="Age at follow-up", ylab="T1D: mortality per 1000 PY (2015-01-01)",
+       xlim=c(10,90), ylim=c(0.2,200) )
> abline( v=e.pr, col=gray(0.8) )
> abline( h=1 )
> axis( side=1 )
> matlines( nd$Ax, t1pr,
+           type="l", lwd=c(3,1,1), lty=1,
+           col=rep(c("blue","red","gray"),each=3) )
> matlines( nd0$Ax, t0pr,
+           type="l", lwd=c(3,1,1), lty=3,
+           col=rep(c("blue","red","gray"),each=3) )
> axis( side=4, at=c(5,10,20)/10 )
> axis( side=4, at=c(5:15,20)/10, label=NA, tcl=-0.3 )
> mtext( "M/F rate-ratio", side=4, at=1, line=2, las=0 )
```

Then we plot the period-RR, extracting the values from the model object:

```
> pr.P <- seq( 1996, 2017, 0.2 )
> rf.P <- 2015
> CP <- Ns( pr.P, knots=p.kn ) - Ns( rep(rf.P,length(pr.P)), knots=p.kn )
> t1rr <- cbind( ci.exp( mt1, ctr.mat=CP, subset="Px" ),
+                ci.exp( ft1, ctr.mat=CP, subset="Px" ) )
```

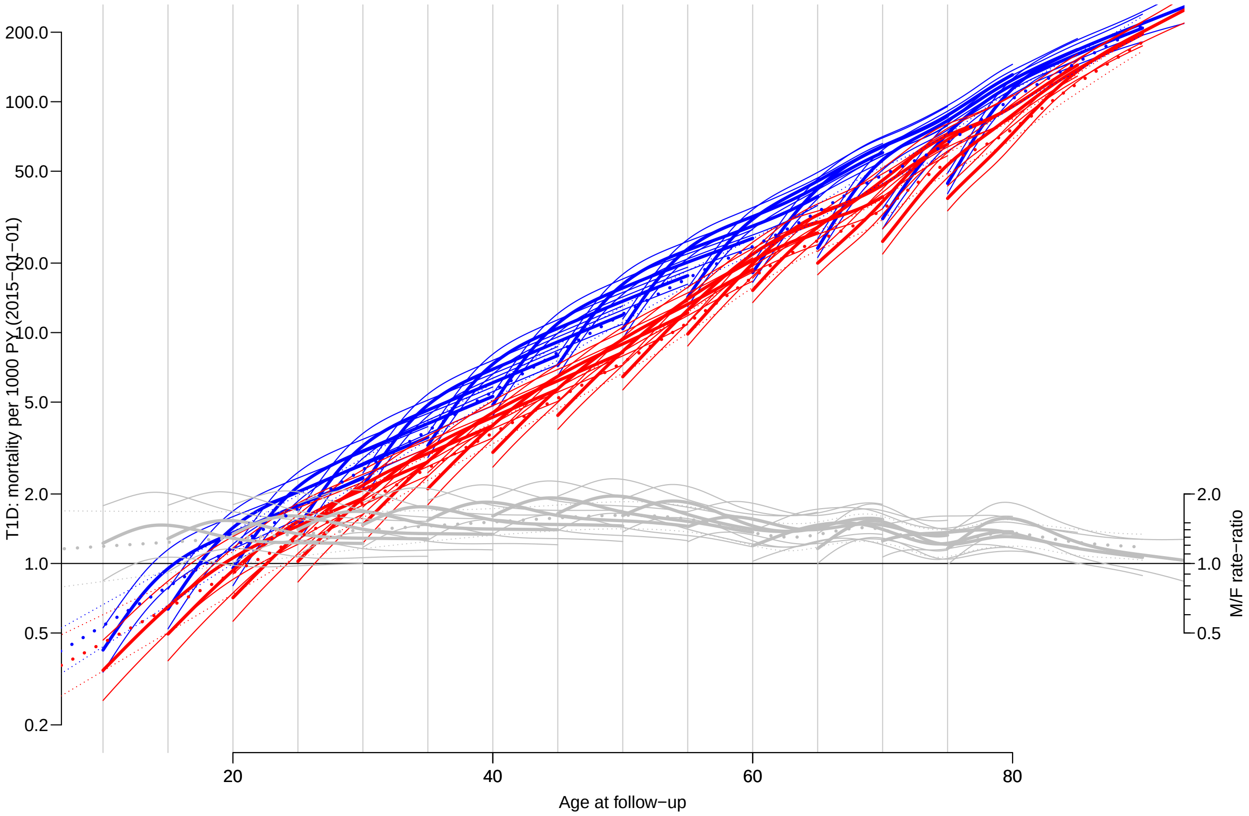

Figure 5.1: *Rates of T1D mortality by age and duration of diabetes. Each curve corresponds to fixed age at diagnosis, strating at the age at diagnosis. Thin lines indicate 95% confidence intervals; red curvers are women, blue men.* `./mort-ainc-t1`

```
> # The following should be doable in Epi 2.27 :
> # ct.lst <-  list( data.frame(Px=pr.P), data.frame(Px=rf.P) )
> # t1rr <- cbind( ci.exp( mt1, ct.lst ), ci.exp( ft1, ct.lst ) )
> par( mar=c(3,3,1,3), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> plot( NA, log="y", xlim=c(1995,2017), ylim=c(0.5,2),
+       xlab="Date of follow-up",
+       ylab="T1D: mortality RR relative to 2015-01-01" )
> abline( h=1 )
> matlines( pr.P, t1rr,
+           type="l", lwd=c(3,1,1), lty=1,
+           col=rep(c("blue","red"),each=3) )
```

The cohort residuals are shown as a check of model-fit:

```
> pr.C <- 1910:1990
> CC <- cbind( 1, Ns( pr.C, knots=c.kn ) )
> t1cr <- cbind( ci.exp( mr1, ctr.mat=CC ),
+                ci.exp( fr1, ctr.mat=CC ) )
> par( mar=c(3,3,1,3), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> plot( NA, log="y", xlim=c(1910,1990), ylim=c(0.5,2), #c(0.8,200)/10,
+       xlab="Date of birth",
+       ylab="T1D: mortality RR residuals" )
> abline( h=1 )
> matlines( pr.C, t1cr, type="l", lwd=c(3,1,1), lty=1,
+                       col=rep(c("blue","red"),each=3) )
```

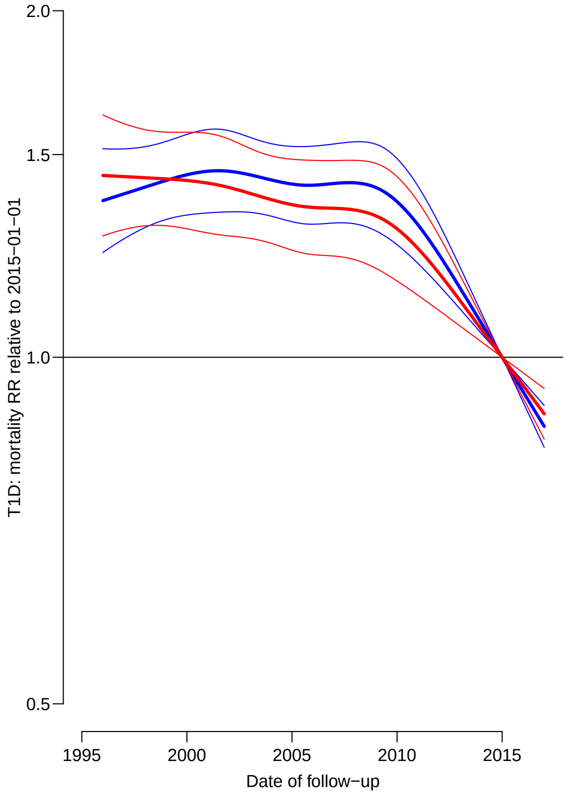

Figure 5.2: *Rate-ratio of T1D mortality reletive to 2015-01-01. Thin lines indicate 95% confidence intervals; red curves are women, blue men.* `./mort-prr-t1`

From figure 5.3 we see that there is no residual confounding by date of birth – the deviations are very small.

5.2 Automation

This analysis will be done for T2 patients as well as for T1 and T2 only diagnosed after 1995-12-31. Hence wit will be suitable to automate it all in a single function that returns the estimated rates and RRs in a list.

```
> mort.res <-
+ function( rt1 )
+     {
+ # Knots for the splines
+ a.kn <-      with( rt1, quantile( rep(Ax    ,D), (1:8-0.5)/8 ) )
+ d.kn <- c(0,with( rt1, quantile( rep(   dur,D), (1:4-0.0)/5 ) ))
+ e.kn <-      with( rt1, quantile( rep(Ax-dur,D), (1:5-0.5)/5 ) )
+ p.kn <-      with( rt1, quantile( rep(   Px ,D), (1:5-0.5)/5 ) )
+ c.kn <-      with( rt1, quantile( rep(Px-Ax ,D), (1:9-0.5)/9 ) )
+
+ # Models for mortality
+ mt0 <- glm( D ~ Ns( Ax    , knots=a.kn ) +
+                 Ns( Px    , knots=p.kn ),
+             offset = log(Y),
+             family = poisson,
```

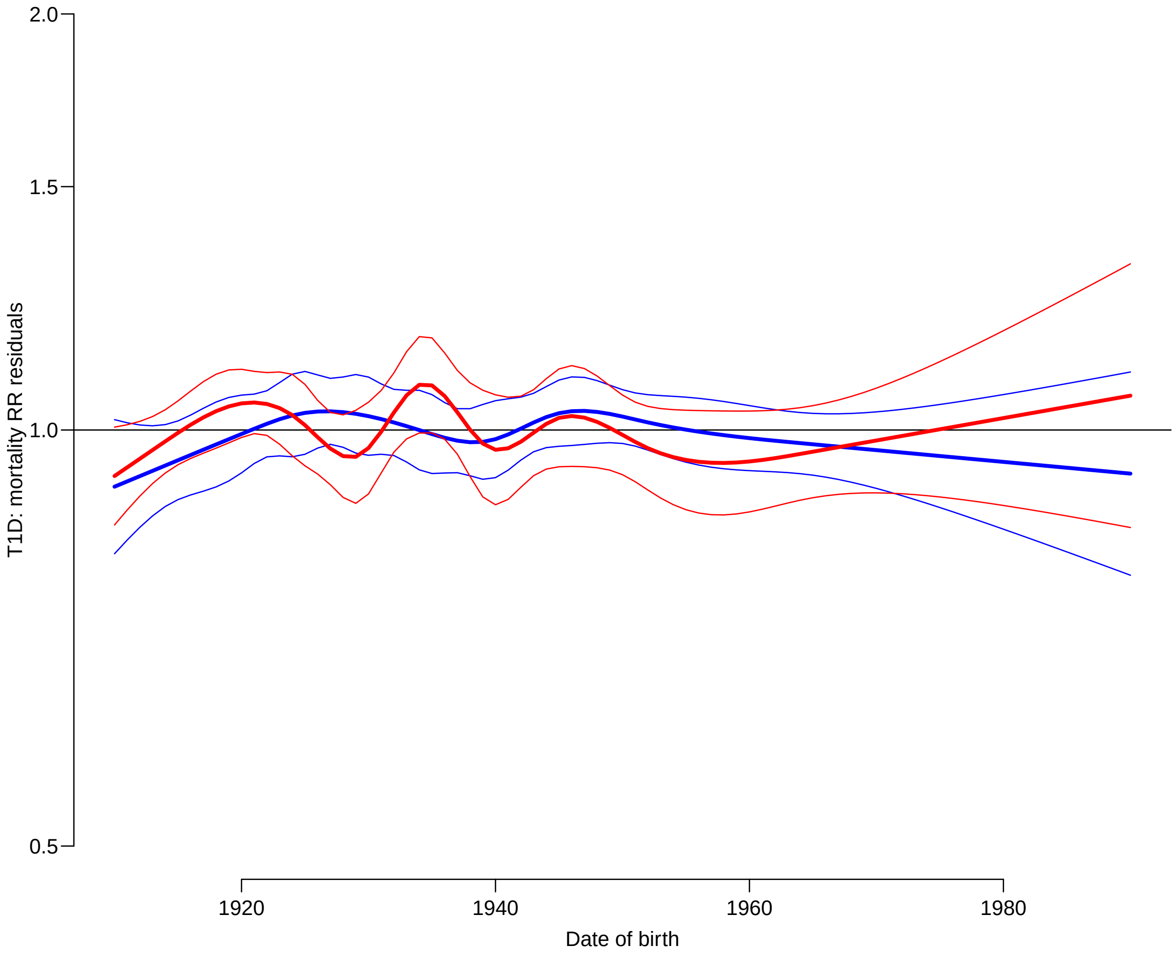

Figure 5.3: *T1D mortality residuals by date of birth. Thin lines indicate 95% confidence intervals; red curves are women, blue men.* `./mort-crr-t1`

```
+                data = subset( rt1, sex=="M" ) )
+ ft0 <- update( mt0, data = subset( rt1, sex=="F" ) )
+ mt1 <- glm( D ~ Ns( Ax     , knots=a.kn ) +
+                 Ns(     dur, knots=d.kn ) +
+                 Ns( Ax-dur, knots=e.kn ) +
+                 Ns( Px     , knots=p.kn ),
+              offset = log(Y),
+              family = poisson,
+                data = subset( rt1, sex=="M" ) )
+ ft1 <- update( mt1, data = subset( rt1, sex=="F" ) )
+ ml1 <- update( mt1, . ~ . - Ns( Px, knots=p.kn ) + Px )
+ fl1 <- update( ft1, . ~ . - Ns( Px, knots=p.kn ) + Px )
+
+ # Summary of period effect
+ test.tr <- matrix( NA, 2, 5 )
+ rownames( test.tr ) <- c("M","F")
+ colnames( test.tr ) <- c("P(lin)","P(sl)","Chg %/y","lo","hi")
+ test.tr["M","P(lin)"] <- anova( mt1, ml1, test="Chisq" )[2,5]
+ test.tr["F","P(lin)"] <- anova( ft1, fl1, test="Chisq" )[2,5]
+ test.tr["M","P(sl)"] <- ci.lin( ml1, subset="Px" )[,"P"]
```

```
+ test.tr["F","P(sl)"] <- ci.lin( fl1, subset="Px" )[,"P"]
+ test.tr["M",c(3,5,4)] <- (ci.exp( ml1, subset="Px" ) - 1)*100
+ test.tr["F",c(3,5,4)] <- (ci.exp( fl1, subset="Px" ) - 1)*100
+
+ # Residuals by cohort
+ mr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(predict(mt1,type="response")),
+             family = poisson,
+               data = subset(rt1,sex=="M") )
+ fr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(predict(ft1,type="response")),
+             family = poisson,
+               data = subset(rt1,sex=="F") )
+
+ # Prediction data frames
+ e.pr <- seq(10,75,5)
+ d.pr <- c( NA, seq(0,20,0.1) )
+ nd0 <- data.frame( Ax=seq( 5,90,0.2), Px=2015, Y=1 )
+ nd <- data.frame( expand.grid( d.pr, e.pr ) )
+ names( nd ) <- c("dur","Ae")
+ nd <- transform( nd, Ax = Ae+dur,
+                      Px = 2015,
+                       Y = 1 )
+
+ # Age-specific rates
+ t0pr <- cbind( ci.pred( mt0, nd0 ), ci.pred( ft0, nd0 ) )
+ t0pr <- cbind( t0pr, ci.ratio(t0pr[,1:3],t0pr[,1:3+3]) )
+ t1pr <- cbind( ci.pred( mt1, nd ), ci.pred( ft1, nd ) )
+ t1pr <- cbind( t1pr, ci.ratio(t1pr[,1:3],t1pr[,1:3+3]) )
+
+ # Period-specific RR
+ pr.P <- seq( 1996, 2017, 0.2 )
+ rf.P <- 2015
+ CP <- Ns( pr.P, knots=p.kn ) - Ns( rep(rf.P,length(pr.P)), knots=p.kn )
+ t1rr <- cbind( ci.exp( mt1, ctr.mat=CP, subset="Px" ),
+                ci.exp( ft1, ctr.mat=CP, subset="Px" ) )
+
+ # Cohort residuals
+ pr.C <- 1910:1990
+ CC <- cbind( 1, Ns( pr.C, knots=c.kn ) )
+ t1cr <- cbind( ci.exp( mr1, ctr.mat=CC ),
+                ci.exp( fr1, ctr.mat=CC ) )
+
+ # return results
+ list( tr = test.tr,
+       A0 = data.frame(nd0[,"Ax"],t0pr),
+       Ad = data.frame(nd[,c("Ae","dur","Ax")],t1pr),
+       Pr = data.frame(P=pr.P,t1rr),
+       Cr = data.frame(C=pr.C,t1cr) )
+ }
> rT1 <- mort.res( subset( rt, state=="T1" ) )
> rT2 <- mort.res( subset( rt, state=="T2" ) )
> rDM <- mort.res( subset( rt, state%in%c("T1","T2") ) )
> LT1 <- mort.res( subset( rtL, state=="T1" ) )
> LT2 <- mort.res( subset( rtL, state=="T2" ) )
> LDM <- mort.res( subset( rtL, state %in% c("T1","T2") ) )
> save( rT1, rT2, rDM,
```

```
+       LT1, LT2, LDM, file="../data/mort-res.Rda" )
```

5.3 Plotting mortality

```
> load( file="../data/mort-res.Rda" )
> lls()
```

```
   name     mode     class            dim             size(Kb)
1  a.kn     numeric  numeric          8                    0.7
2  c.kn     numeric  numeric          9                    0.9
3  CC       numeric  matrix           81 9                 6.6
4  CP       numeric  cns basis matrix 106 4                4.9
5  d.kn     numeric  numeric          4                    0.4
6  d.pr     numeric  numeric          202                  1.6
7  e.kn     numeric  numeric          5                    0.5
8  e.pr     numeric  numeric          14                   0.2
9  fl1      list     glm lm           30              79,845.6
10 fr1      list     glm lm           30              74,460.0
11 ft0      list     glm lm           30              73,698.7
12 ft1      list     glm lm           30              84,451.9
13 LDM      list     list             5                  493.9
14 LT1      list     list             5                  493.9
15 LT2      list     list             5                  493.9
16 ml1      list     glm lm           30              77,417.9
17 mort.res function function         1                  603.7
18 mr1      list     glm lm           30              72,195.7
19 mt0      list     glm lm           30              71,457.7
20 mt1      list     glm lm           30              81,884.1
21 nd       list     data.frame       2828 5             111.5
22 nd0      list     data.frame       426 3               10.9
23 p.kn     numeric  numeric          5                    0.5
24 pr.C     numeric  integer          81                   0.4
25 pr.P     numeric  numeric          106                  0.9
26 rDM      list     list             5                  493.9
27 rf.P     numeric  numeric          1                    0.0
28 rt       list     data.frame       384545 13       36,054.0
29 rt1      list     data.frame       193282 13       18,878.1
30 rT1      list     list             5                  493.9
31 rT2      list     list             5                  493.9
32 rtL      list     data.frame       188843 13       17,706.9
33 t0pr     numeric  matrix           426 9               54.0
34 t1cr     numeric  matrix           81 6                 4.4
35 t1pr     numeric  matrix           2828 9             354.2
36 t1rr     numeric  matrix           106 6                5.6
```

```
> str( LDM )
```

```
List of 5
 $ tr: num [1:2, 1:5] 7.97e-02 1.25e-08 1.17e-305 6.09e-134 -3.46 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : chr [1:2] "M" "F"
  .. ..$ : chr [1:5] "P(lin)" "P(sl)" "Chg %/y" "lo" ...
 $ A0:'data.frame':       426 obs. of  10 variables:
  ..$ nd0....Ax..: num [1:426] 5 5.2 5.4 5.6 5.8 6 6.2 6.4 6.6 6.8 ...
  ..$ Estimate   : num [1:426] 0.478 0.484 0.491 0.497 0.504 ...
  ..$ X2.5.      : num [1:426] 0.415 0.421 0.427 0.433 0.439 ...
```

```
  ..$ X97.5.     : num [1:426] 0.55 0.557 0.564 0.571 0.579 ...
  ..$ Estimate.1 : num [1:426] 0.316 0.32 0.324 0.329 0.333 ...
  ..$ X2.5..1    : num [1:426] 0.262 0.266 0.269 0.273 0.277 ...
  ..$ X97.5..1   : num [1:426] 0.38 0.385 0.39 0.395 0.4 ...
  ..$ Estimate.2 : num [1:426] 1.52 1.51 1.51 1.51 1.51 ...
  ..$ X2.5..2    : num [1:426] 1.2 1.2 1.2 1.2 1.2 ...
  ..$ X97.5..2   : num [1:426] 1.91 1.91 1.91 1.91 1.9 ...
 $ Ad:'data.frame':        2828 obs. of  12 variables:
  ..$ Ae        : num [1:2828] 10 10 10 10 10 10 10 10 10 10 ...
  ..$ dur       : num [1:2828] NA 0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 ...
  ..$ Ax        : num [1:2828] NA 10 10.1 10.2 10.3 10.4 10.5 10.6 10.7 10.8 ...
  ..$ Estimate  : num [1:2828] NA 0.897 0.87 0.844 0.819 ...
  ..$ X2.5.     : num [1:2828] NA 0.788 0.765 0.742 0.721 ...
  ..$ X97.5.    : num [1:2828] NA 1.022 0.991 0.96 0.931 ...
  ..$ Estimate.1: num [1:2828] NA 0.687 0.66 0.634 0.61 ...
  ..$ X2.5..1   : num [1:2828] NA 0.578 0.556 0.534 0.514 ...
  ..$ X97.5..1  : num [1:2828] NA 0.817 0.784 0.753 0.724 ...
  ..$ Estimate.2: num [1:2828] NA 1.31 1.32 1.33 1.34 ...
  ..$ X2.5..2   : num [1:2828] NA 1.05 1.06 1.07 1.08 ...
  ..$ X97.5..2  : num [1:2828] NA 1.62 1.64 1.65 1.66 ...
 $ Pr:'data.frame':        106 obs. of  7 variables:
  ..$ P          : num [1:106] 1996 1996 1996 1997 1997 ...
  ..$ exp.Est..  : num [1:106] 1.89 1.88 1.87 1.85 1.84 ...
  ..$ X2.5.      : num [1:106] 1.8 1.79 1.78 1.77 1.76 ...
  ..$ X97.5.     : num [1:106] 1.99 1.97 1.96 1.94 1.93 ...
  ..$ exp.Est...1: num [1:106] 1.54 1.53 1.52 1.52 1.51 ...
  ..$ X2.5..1    : num [1:106] 1.45 1.45 1.44 1.44 1.44 ...
  ..$ X97.5..1   : num [1:106] 1.63 1.62 1.61 1.6 1.59 ...
 $ Cr:'data.frame':        81 obs. of  7 variables:
  ..$ C          : int [1:81] 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 ...
  ..$ exp.Est..  : num [1:81] 0.891 0.906 0.922 0.938 0.955 ...
  ..$ X2.5.      : num [1:81] 0.849 0.869 0.89 0.91 0.929 ...
  ..$ X97.5.     : num [1:81] 0.934 0.945 0.956 0.968 0.981 ...
  ..$ exp.Est...1: num [1:81] 0.967 0.973 0.978 0.984 0.99 ...
  ..$ X2.5..1    : num [1:81] 0.936 0.945 0.954 0.962 0.97 ...
  ..$ X97.5..1   : num [1:81] 0.999 1.001 1.003 1.006 1.01 ...
```

We will make two sets of 3 × 2 plots; one based on the total dataset, ignoring problems with the date of diagnosis; and one which is only based on the persons diagnosed after 1996-01-01, thus with a better precision of date of diagnosis.

```
> pltp <-
+ function( rT1, coh=FALSE, e.pr=seq(10,70,10), lab="" )
+     {
+ rdf <- subset( rT1$Ad, Ae %in% e.pr )
+ art <- rT1$A0
+ prt <- rT1$Pr
+ crt <- rT1$Cr
+
+ plot( NA, log="y", xlim=c(10,90), ylim=c(0.2,200),
+       xlab="", ylab="" )
+ axis( side=1, at=1:9*10, labels=NA, tcl=-0.5 )
+ axis( side=1, at=2:18*5, labels=NA, tcl=-0.3 )
+ abline( v=e.pr, col=gray(0.8) )
+ abline( h=1 )
+ matlines( rdf[,"Ax"], rdf[,-c(1:3)],
+           type="l", lwd=c(3,1,1), lty=1,
```

```
+            col=rep(c("blue","red","gray"),each=3) )
+ matlines( art[,1], art[,-1],
+            type="l", lwd=c(3,1,1), lty=3,
+            col=rep(c("blue","red","gray"),each=3) )
+ axis( side=4, at=c(5,10,20)/10 )
+ axis( side=4, at=c(5:15,20)/10, label=NA, tcl=-0.3 )
+ text( 10, 200, lab, font=2, adj=0, cex=1.5 )
+
+ plot( NA, log="y", xlim=c(1995,2017), ylim=c(0.5,2.5),
+       xlab="",
+       ylab="" )
+ abline( h=1 )
+ axis( side=1, at=1996:2017, labels=NA, tcl=-0.3 )
+ matlines( prt$P, prt[,-1],
+            type="l", lwd=c(3,1,1), lty=1,
+            col=rep(c("blue","red"),each=3) )
+
+ if( coh ){
+ plot( NA, log="y", xlim=c(1910,1990), ylim=c(0.5,2.5),
+       xlab="",
+       ylab="" )
+ abline( h=1 )
+ matlines( crt$C, crt[,-1],
+            type="l", lwd=c(3,1,1), lty=1,
+            col=rep(c("blue","red"),each=3) )
+          }
+      }
> layout( matrix(1:9,3,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rT1, coh=T, lab="T1D" )
> pltp( rT2, coh=T, lab="T2D" )
> pltp( rDM, coh=T, lab="DM" )
> mtext( c("Mortality rate per 1000 py",
+          "Rate ratio relative to 2015-01-01",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up",
+          "Date of birth"),
+        at = c(4,10,14)/16,
+        side=1, outer=TRUE, cex=0.67 )
```

```
> layout( matrix(1:9,3,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( LT1, coh=T )
> pltp( LT2, coh=T )
> pltp( LDM, coh=T )
> mtext( c("Mortality rate per 1000 py",
+          "Rate ratio relative to 2015-01-01",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up",
+          "Date of birth"),
+        at = c(4,10,14)/16,
+        side=1, outer=TRUE, cex=0.67 )
```

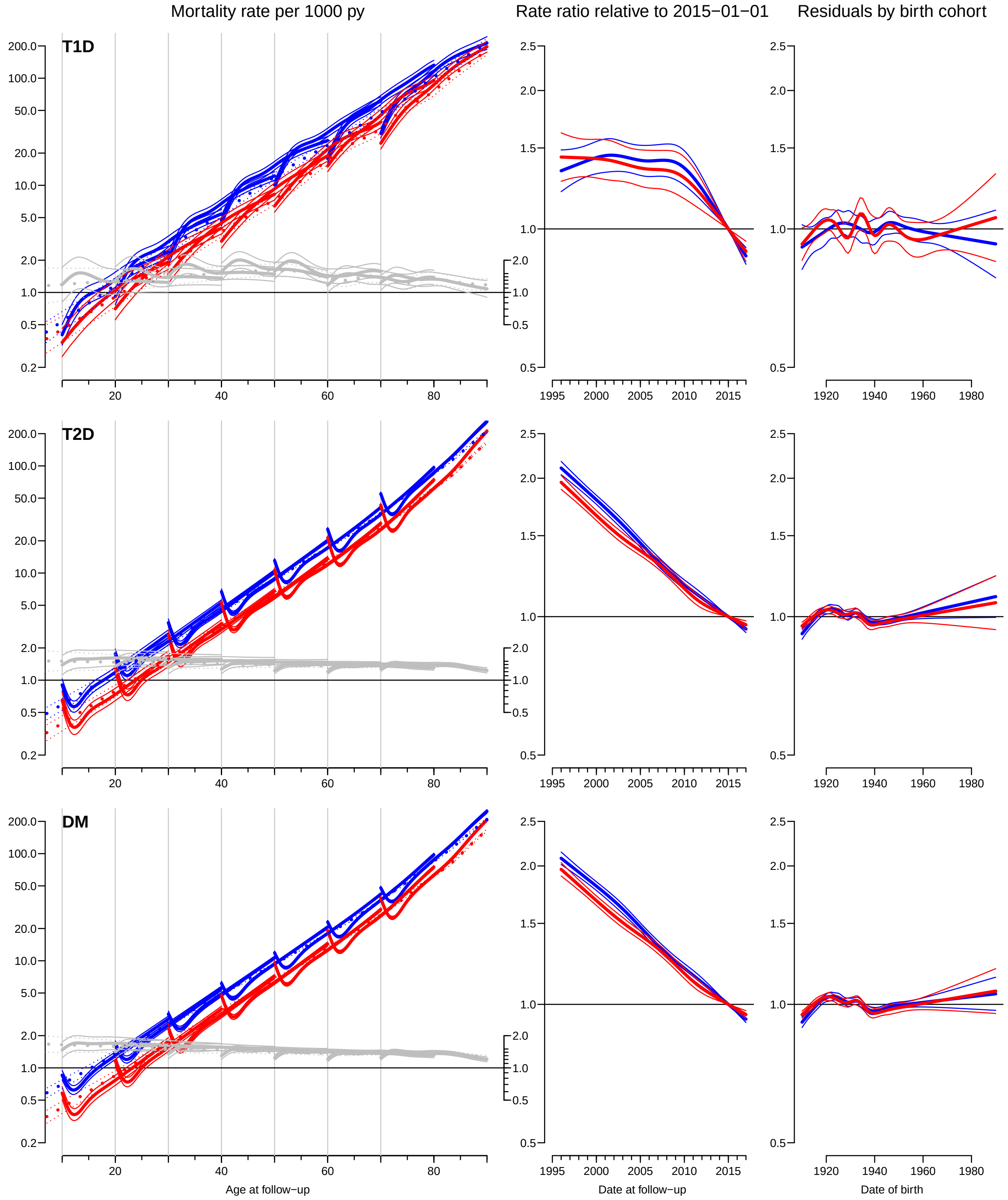

Figure 5.4: *Mortality, HR relative to 2015-01-01 and birth cohort residuals. Leftmost plot shows the mortality rates at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are mortality rates modeled ignoring age at diagnosis and duration of diabetes.*
Plots are based on all persons in DMreg, assuming that entry dates before 1996-01-01 are usable.
Red curves are for women, blue for me. `./mort-all-m`

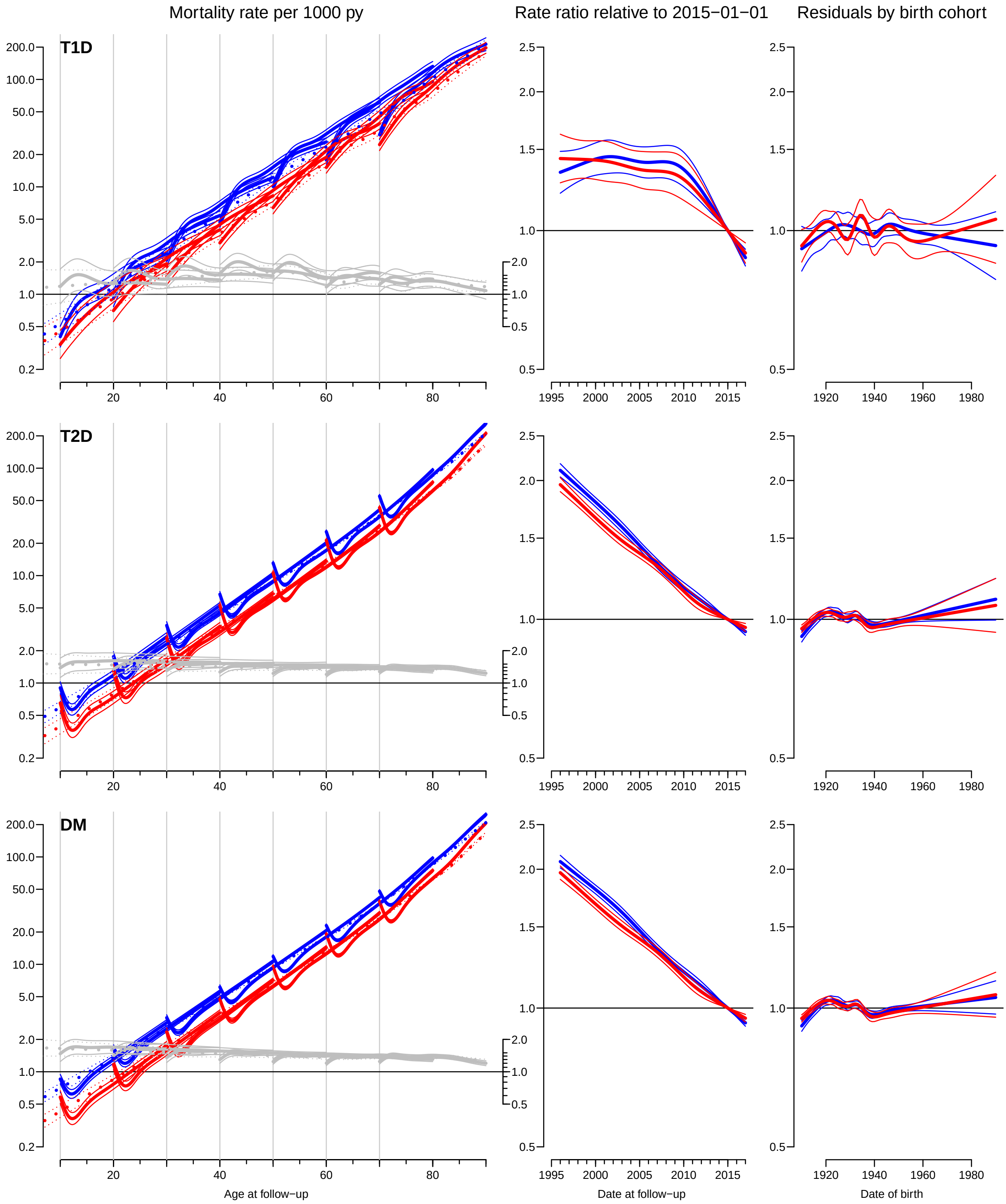

Figure 5.5: *Mortality, HR relative to 2015-01-01 and birth cohort residuals. Leftmost plot shows the mortality rates at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration.*
Plots are based only on persons in DMreg with date of diagnosis afte 1996-01-10, assuming that entry dates before 1996-01-01 are useless.
Red curves are for women, blue for me. `./mort-all-m`

Bibliography