Analyses based on the reconstructed Danish Diabetes Register

SDC
May 2019
http://bendixcarstensen.com/DMreg/NewAna.pdf
Draft version 15

Compiled Monday 27th May, 2019, 17:15
from: /home/bendix/sdc/DMreg/NewReg/nyr/NewAna.tex

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

Contents

1 Data 1

2 Prerequisites and population forecast 2
- 2.1 Data structures used and created . . . 2
- 2.2 Prerequisites . . . 4
  - 2.2.1 Points . . . 4
  - 2.2.2 Knots . . . 5
  - 2.2.3 Cautious printing . . . 5
  - 2.2.4 Nice printing . . . 6
- 2.3 DST population forecasts . . . 6

3 Register overview — first contact 9
- 3.1 The register data . . . 9
- 3.2 Overview of the register . . . 10
- 3.3 Inclusion criteria . . . 13
  - 3.3.1 Recorded inclusion dates . . . 13
  - 3.3.2 Inclusion criteria by year . . . 14
- 3.4 Meeting more than one criterion . . . 22
  - 3.4.1 Defining the follow-up . . . 22
- 3.5 Meeting more than one criterion . . . 24
  - 3.5.1 Defining the follow-up . . . 24
  - 3.5.2 Analysis of rate of meeting 2nd criterion . . . 26

4 Diabetes prevalence 31
- 4.1 Prevalence data . . . 31
  - 4.1.1 Raw tables . . . 32
  - 4.1.2 Analysis dataset . . . 33
- 4.2 Age-distribution of prevalent diabetes cases . . . 35
  - 4.2.1 Population pyramids . . . 37
- 4.3 Models for prevalence . . . 38
  - 4.3.1 Fration of patients with T1D . . . 50
- 4.4 Consistency of population size . . . 54

5 Rate data 57

6 DM incidence **61**
6.1 Age-Period-Cohort models . . . 64
6.1.1 Time-trends in rates . . . 67
6.2 Model for incidence rates . . . 71
6.2.1 `gam` models . . . 72
6.2.2 Local drifts . . . 72
6.3 Collecting estimated incidence rates . . . 73
6.3.1 Rates by age for different periods (cross-sectional) . . . 74
6.3.2 Rates by calendar time in different ages . . . 79
6.4 Incidence summary . . . 83
6.4.1 Saving the fitted models . . . 83

7 Mortality **84**
7.1 Mortality overview . . . 85
7.2 Type 1 diabetes patient mortality . . . 88
7.3 Mortality for T1, T2 and all DM . . . 94
7.4 Plotting mortality rates . . . 96
7.5 Comparison of T1 and T2 mortality . . . 102
7.5.1 Direct comparison of T1D vs. T2D mortality . . . 105
Trend in T1 / T2 mortality rate ratio . . . 105
7.6 SMR . . . 106
7.6.1 Plotting the SMR . . . 108

8 Mortality after 2005 **113**
8.1 Mortality overview . . . 114
8.2 Type 1 diabetes patient mortality . . . 116
8.3 Mortality for T1, T2 and all DM . . . 120
8.4 Plotting mortality rates . . . 122
8.5 Comparison of T1 and T2 mortality . . . 129
8.5.1 Direct comparison of T1D vs. T2D mortality . . . 131
8.6 SMR . . . 131
8.6.1 Plotting the SMR . . . 133

9 Analysis and prediction of rates **138**
9.1 Analysis data . . . 138
9.1.1 Arrays of predictions . . . 142
9.1.2 Datasets for rate modeling . . . 143
9.2 Models for incidence and mortality rates . . . 144
9.3 Incidence rates . . . 144
9.3.1 Early and late trends in incidence rates . . . 145
9.3.2 Incidence rate predictions . . . 146
9.3.3 Calendar time trends . . . 147
9.4 Attenuated extrapolation . . . 148
9.4.1 Theory . . . 152
Exponential damping of a drift . . . 152
Implementation in APC-models . . . 153
9.4.2 A discrete time shortcut . . . 154

Adding a drift . . . 155
Implementation . . . 156
9.5 Mortality rates . . . 159
9.5.1 Diabetes patients . . . 159
9.5.2 Persons without diabetes . . . 160
9.5.3 Relative mortality (SMR) . . . 161
9.6 Average trends . . . 163
9.7 Time trends in rates . . . 163
9.7.1 Age-specific rates by date . . . 164
Illustrative plots . . . 164
9.7.2 Date specific rates by age . . . 166
9.8 Extrapolation of rates . . . 167
Digression on `apply` . . . 167
9.8.1 Attenuated extrapolations . . . 168
9.9 The projected rates beyond 2016 . . . 169

10 Predicting prevalence of diabetes **179**
10.1 Predicted rates . . . 179
10.2 Transition probabilities . . . 182
10.3 Prediction of the observed prevalences . . . 183
10.4 The actual numbers of diabetes patients in Denmark . . . 185
10.5 Time trends in prevalent number of DM patients . . . 189
10.5.1 Analyses by sex . . . 199
10.6 Time trends in old age patients . . . 208

11 Components of prevalence **213**
11.1 Transition probabilities . . . 214
11.2 Prediction of the observed prevalences . . . 216
11.2.1 Checking the prediction . . . 219
11.3 How much is attributable to what? . . . 223
11.4 Number of diabetes patients in Denmark . . . 226
11.5 Time trends in the components . . . 231

12 Lifetime lost to diabetes **236**
12.1 Theory . . . 236
12.1.1 Overall mortality (any cause) . . . 236
12.1.2 Years of life lost by cause of death . . . 237
12.1.3 Multistate set-up for register data . . . 237
12.1.4 Practical implementation of multistate model . . . 239
12.2 Models for rates . . . 240
12.2.1 Modeling the transition rates . . . 243
Rates by age at different dates . . . 246
Rates by date for different ages . . . 247
12.2.2 Checking with all-cause mortality . . . 249
12.2.3 Comparing mortality rates . . . 256
12.3 Transition probabilities . . . 258
12.3.1 Set-up for YLL . . . 258

12.4 State occupancy probabilities . . . 261
12.5 State probabilities for 2016 . . . 262
12.5.1 State probabilities from age 50 . . . 263
12.6 Lifetime lost . . . 266
12.6.1 Lifetimne lost by cause of death . . . 275
12.6.2 Total lifetime lost to diabetes . . . 276
12.7 Population burden of diabetes . . . 285
12.7.1 Diabetes population . . . 285
Prevalent cases . . . 285
Incident cases . . . 286
12.7.2 Total fututre lifetime lost . . . 287
Prevalent cases . . . 287
Incident cases . . . 292
12.8 Life time risk . . . 298
12.9 Transition probabilities for life time risk . . . 301
12.10 State occupancy probabilities . . . 303
12.11 Cumulative risk of T2D . . . 305
12.12 Lifetime risk of diabetes . . . 307
12.12.1 Sojourn times . . . 309

References **317**

Chapter 1

Data

The maintenance of the National Diabetes Register (NDR) has been discontinued by the Health Data Authority (Sundhedsdatastyrelsen). It has been replaced by the Register of Selected Chronic Diseases (RUKS — Register for Udvalgte Kroniske Sygdomme) which however does not encompass precisely the same persons.

A replacement of the NDR with greater precision than both RUKS and NDR has therefore beeen constructed; it is documented in the report
http://BendixCarstensen.com/DMreg/NewReg.pdf, which also documents the construction of the follow-up and prevalence data used in the analyses done here.

. . . now input from `init.tex`

```
> library( Epi )
> library( splines )
> start()
Home folder E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-01-01 12:10:51

R version 3.5.0 (2018-04-23)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows Server 2012 R2 x64 (build 9600)

Matrix products: default

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

other attached packages:
[1] Epi_2.30

loaded via a namespace (and not attached):
 [1] cmprsk_2.2-7      zoo_1.8-1         MASS_7.3-49       compiler_3.5.0
 [5] Matrix_1.2-14     plyr_1.8.4        parallel_3.5.0    tools_3.5.0
 [9] survival_2.41-3   etm_0.6-2         Rcpp_0.12.16      grid_3.5.0
[13] numDeriv_2016.8-1 lattice_0.20-35
```

Chapter 2

Prerequisites and population forecast

This document contains a number of different analyses, all based on either the register file itself, the prevalence dataset or the rates dataset derived from the SAS-files underlying the register construction.

Some of the analyses produce results in the form of estimated prevalences or rates that are used in subsequent analyses. Below is a list of `.Rda` files produced in each chapter, their content and later use, which can be used as a quick look-up of what is loaded and the content of the data structures.

2.1 Data structures used and created

All data files are put in the folder `..\data`.

Prerequisites saves the R-files:

`inits.Rda` , with the objects:

- `qn` function to return equally spaced probabilities
- `fc, fCp, fCtable` functions to print numbers and (`f`)tables with comma separator of numbers
- `cstr` abbreviated versions of `str` that does not list single items
- `int` Scalar, interval length ti be used for prediction calculations in predictions and components of prevalence
- `a.pt, t.pt, p.pt, d.pt` vectors; midpoints of age (`a.pt`, 0–100), period (`p.pt`, 1996–2040) and duration (`d.pt`, 0–20) intervals, while `t.pt` is the endpoints of intervals from 1996 through 2040.
- `nk.a, nk.p, nk.c, nk.d`, scalars; number of knots to be used for the natural splines for age, period, cohort and duration of DM.

`pop` objects with projected population size (from Statistics Denmark) at 1 January 2017–2040:

- `beff` data frame with variables `sex` (M,W), `A` (0–100), `P` (2017–2040) and `N`, population size
- `befr` the same but subdivided by region of residence (5 levels).

Prevalence saves two R-files:

`prevN` with object

- `prN` data frame classified by `reg` (region) `sex` (M/W) `A` (0–99) `P` (1996–2017) with variables `T1`, `T2`, `DM` (all diabetes), `nD` (non-diabetes) and `N` (total population)

`prevalences` with objects

- `parr` array of predicted prevalences ($\in [0,1]$) classified by `mod` model (`glm`, `gam`), `typ` outcome (T1, T2, DM), `sex` (M,W), `A` age (0 to 100 by 0.5) `T` date (1996–2017) `Est, lo, hi`.
- `mod`: List of `glm` models for prevalence at 2017-01-01 with dimensions `typ` (T1, T2, DM) and `sex` (M, W)
- `akn`, as `mod` but with the knots used in the natural splines in `mod` — this is needed if models is to be used for predictions
- `pr.obs` array of observed *number* of prevalent cases as of 1996-01-01 – 2017-01-01 classified by `typ` (T1, T2, DM) and `sex` (M, W), and 1-year classes of age.
- `pr.ini, pr.fin`, arrays of predicted prevalences at 1996-01-01, resp. 2017-01-01 classified by `typ` (T1, T2, DM) and `sex` (M, W), and evaluated at the midpoint of 1200 one-month age-intervals from age 0–100.

Rates reads the SAS-file `FUtot` and saves the R-file:

`rt` with objects

- `rt`, `rtL` data frames of follow-up in the entire population in 1-year Lexis triangles by age (`Ax`) and calendar time (`Px`) classified also by diabetes duration (`dur`, (0 for noDM, midpoint of interval for T1 and T2), `state` (noDM, T1, T2) and `sex` (M,W).
  `rt` has all follow-up time, `rtL` excludes follow-up among T1 and T2 with date of diagnosis before 1996-01-01, so the dataset with reliable duration information.

Incidence saves no files of interest; loads the files `inits` and `rt` saved in chapters **init** resp **rates**.

Mortality saves two files:

`mort-res` with three objects of similar structure

- `rT1`, `rT2` and `rDM`; referring to mortality rates of T1, T2 and all DM; each one a list of 5 objects:
  - `rt` Matrix classified by sex (M,W) and test-p-values for linearity, slope and estimate of slope with c.i. from age-period models with splines, resp.linear effects of period.
  - `A0` Data frame with ages (5–90), mortality rates for M, W and M/W RR at 2015-01-01 from an age-period model; 9 columns based on patients diagnosed after 1996-01-01 followed by 9 columns based on all patients — a total of $1 + 9 + 9 = 19$ columns

  - `Ad` Data frame with age at diagnosis (10, 15,. . . ,75), duration (0.1, 0.2,. . . ,20) and current age, and mortality for M, W and M/W RR as of 2015-01-01 — a total of $3 + 9 = 12$ columns. Based on a model with age at diagnosis, current age, duration and period for patients diagnosed after 1996-01-01.
  - `Pr` Data frame of period (1996,1996.2,. . . ,2017) and RR relative to 2015 for M, W from the above model.
  - `Cr` Data frame of cohort (1910,1911,. . . ,1990) and residual RR from the above model fitted cohort effects from from a cohort-only model using log-fitted values as offset.

`smr-res` contains the same as `mort-res`, except that the components are called `xT1`, `xT2` and `xDM`, and contain analyses of SMR instead of mortality rates.

Analysis and prediction Saves two R-files:

`rate-dat` holds data frames

  - `incdat` (incidence of all DM),
  - `mnDdat` (mortality in non-DM persons) and
  - `mDMdat` (mortality among all DM patients, all with variables `A` (age) `P` (period) `C` (cohort) `D` (events) `Y` (person-time), each record representing a Lexis-triangle in the grid 0–100 by 1996–2017 Lexis diagram.

`allrates` holds 3 arrays:

  - `Lambda` (incidence rates of DM),
  - `Mu.nD` (mortality rates for non-DM) and
  - `Mu.DM` (mortality among DM patients. Classified by age (`a`) (midpoints of 1-month intervals 0-100 years), period (`p`) (midpoints of 1-month intervals 1996–2040), sex and model used for prediction beyond 2017 (`ap`, `apc`, `gam`, `LCa`, `att`, `fix`, `p20`, `p40`, `p60`). To be used in future prediction of prevalent cases and prediction of components of DM prevalence.

2.2 Prerequisites

2.2.1 Points

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 2040, although actual reporting will only be till 2030 — mainly 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
> d.pt <- seq(0,20,int)[-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.

2.2.2 Knots

We will use natural splines to model the effects of age, period, cohort and for mortality rates also duration, 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
> nk.d <- 6
```

For the practical location of the spline knots we 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
```

This function generates `nk` equidistant points in the interval (0,1) where the outer points are $1/(2 \times \texttt{nk})$ from the end. The second parameter, `bd`, is for modifying the multiplier 2, so that the outer intervals are $1/(\texttt{nk} \times \texttt{bd})$

We will however also use `gam` from the `mgcv` package to model the non-linear effects.

2.2.3 Cautious printing

A small feature to avoid printing individual data in output:

```
> cstr <- function( reg )
+ {
+ if( !inherits( reg, "data.frame" ) ) stop( "Argument must be a data.frame" )
+ df <- data.frame( no = 1:ncol( reg),
+                   class = sapply( lapply( reg, class ), paste, collapse=", " ) )
+ colnames( df ) <- NULL
+ cat( class(reg), paste( dim( reg ), collapse=" by " ) )
+ print( df )
+ }
```

2.2.4 Nice printing

Finally, we define a couple of functions to enhance readability of large 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, cstr,
+       int, a.pt, t.pt, p.pt, d.pt,
+         nk.a, nk.c, nk.p, nk.d,
+       file="../nydata/inits.Rda" )
```

2.3 DST population forecasts

A prerequisite of the prediction of the prevalent *number* of persons is that we have some knowledge or assumptions about the future *total* population size of Denmark, so we acquired population forecasts from the data bank of DST, and put it in an edited `.csv`-file:

```
> bef <- read.csv2( "../data/bef2040.csv", header=FALSE )
> names( bef )[1:3] <- c("sex","reg","A")
> bef$sex <- Relevel( bef$sex, 2:1 )
> str( bef )
'data.frame':        1010 obs. of  27 variables:
 $ sex: Factor w/ 2 levels "W","M": 2 2 2 2 2 2 2 2 2 2 ...
 $ reg: Factor w/ 5 levels "Midt","Nord",..: 3 3 3 3 3 3 3 3 3 3 ...
 $ A  : int  0 1 2 3 4 5 6 7 8 9 ...
 $ V4 : int  11398 10678 10364 10022 10297 10127 10856 10546 10868 10640 ...
 $ V5 : int  11680 11277 10586 10258 9963 10245 10099 10843 10544 10870 ...
 $ V6 : int  11930 11540 11157 10463 10184 9905 10200 10076 10825 10539 ...
 $ V7 : int  12174 11800 11428 11042 10408 10145 9888 10188 10077 10828 ...
 $ V8 : int  12431 12032 11673 11302 10972 10358 10116 9867 10174 10073 ...
 $ V9 : int  12704 12276 11889 11531 11220 10906 10319 10085 9846 10159 ...
 $ V10: int  12989 12538 12123 11739 11443 11146 10858 10283 10059 9831 ...
 $ V11: int  13251 12811 12372 11963 11644 11363 11091 10815 10254 10042 ...
 $ V12: int  13453 13063 12637 12207 11864 11562 11307 11048 10784 10237 ...
 $ V13: int  13594 13258 12880 12464 12102 11778 11502 11261 11015 10767 ...
 $ V14: int  13686 13393 13068 12700 12354 12012 11716 11455 11226 10999 ...
 $ V15: int  13741 13481 13200 12884 12586 12261 11948 11666 11419 11209 ...
 $ V16: int  13768 13534 13285 13011 12766 12490 12193 11896 11628 11401 ...
 $ V17: int  13774 13559 13336 13095 12891 12667 12420 12139 11856 11609 ...
 $ V18: int  13765 13564 13360 13144 12973 12791 12596 12363 12097 11835 ...
 $ V19: int  13745 13554 13362 13164 13017 12868 12715 12536 12319 12074 ...
 $ V20: int  13718 13534 13352 13167 13039 12912 12792 12655 12492 12297 ...
 $ V21: int  13686 13506 13329 13156 13041 12932 12835 12731 12609 12468 ...
 $ V22: int  13649 13473 13301 13132 13028 12934 12854 12773 12684 12585 ...
 $ V23: int  13607 13438 13269 13105 13006 12921 12856 12793 12726 12659 ...
 $ V24: int  13559 13395 13233 13073 12978 12898 12842 12793 12744 12699 ...
 $ V25: int  13503 13346 13190 13035 12944 12870 12819 12779 12744 12717 ...
 $ V26: int  13439 13291 13142 12995 12910 12838 12791 12756 12731 12717 ...
 $ V27: int  13369 13227 13087 12946 12866 12801 12758 12728 12707 12702 ...
```

```
> befr <- reshape( bef, direction = "long",
+                          varying = 4:27,
+                            times = 2017:2040,
+                          timevar = "P",
+                          v.names = "N" )[,1:5]
> rownames( befr ) <- NULL
> beff <- aggregate( befr$N, befr[,c("sex","A","P")], FUN = sum )
> names( beff )[4] <- "N"
```

Omit persons over 100:

```
> beff <- subset( beff, A<100 )
> attr( beff, "Label" ) <- "DK population forecasts as of 2017 - 2040-01-01"
> str( beff )
'data.frame':        4800 obs. of  4 variables:
 $ sex: Factor w/ 2 levels "W","M": 1 2 1 2 1 2 1 2 1 2 ...
 $ A  : int  0 0 1 1 2 2 3 3 4 4 ...
 $ P  : int  2017 2017 2017 2017 2017 2017 2017 2017 2017 2017 ...
 $ N  : int  30054 31970 28889 30364 28451 30069 28302 29674 29187 31101 ...
 - attr(*, "Label")= chr "DK population forecasts as of 2017 - 2040-01-01"
```

```
> summary( beff )
```

```
 sex            A               P              N
 W:2400   Min.   : 0.00   Min.   :2017   Min.   :  106
 M:2400   1st Qu.:24.75   1st Qu.:2023   1st Qu.:29379
          Median :49.50   Median :2028   Median :34293
          Mean   :49.50   Mean   :2028   Mean   :30193
          3rd Qu.:74.25   3rd Qu.:2034   3rd Qu.:37461
          Max.   :99.00   Max.   :2040   Max.   :45246
```

```
> fCtable( addmargins( xtabs( N ~ P + sex, data=beff ), 2), w=10 )
```

```
     sex           W           M         Sum
P
2017       2,887,581   2,859,967   5,747,548
2018       2,903,785   2,879,674   5,783,459
2019       2,918,257   2,896,601   5,814,858
2020       2,931,826   2,911,734   5,843,560
2021       2,944,824   2,925,660   5,870,484
2022       2,957,649   2,938,924   5,896,573
2023       2,970,556   2,951,933   5,922,489
2024       2,983,488   2,964,768   5,948,256
2025       2,996,356   2,977,487   5,973,843
2026       3,009,043   2,989,922   5,998,965
2027       3,021,447   3,002,017   6,023,464
2028       3,033,430   3,013,728   6,047,158
2029       3,044,924   3,024,967   6,069,891
2030       3,055,949   3,035,725   6,091,674
2031       3,066,384   3,046,018   6,112,402
2032       3,076,292   3,055,782   6,132,074
2033       3,085,635   3,065,093   6,150,728
2034       3,094,437   3,073,984   6,168,421
2035       3,102,691   3,082,439   6,185,130
2036       3,110,390   3,090,527   6,200,917
2037       3,117,511   3,098,191   6,215,702
2038       3,124,114   3,105,499   6,229,613
2039       3,130,243   3,112,513   6,242,756
2040       3,135,975   3,119,202   6,255,177
```

Finally, we save the population predictions for future use:

```
> save( beff, befr, file="../nydata/pop.Rda" )
```

```
----------------------
2019-01-01 at 12:10:52
Time elapsed: 00:00:01
----------------------
```

. . . now input from `reg.tex`

```
---------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-01-06 11:55:03
---------------------------------------
```

Chapter 3

Register overview — first contact

3.1 The register data

First we read the register file and define the relevant variabels as factors.

```
> system.time( rg <- read_sas("../nydata/DMreg.sas7bdat") )
   user  system elapsed
   5.50    0.13   11.21
> reg <- as.data.frame( rg )
```

Here we use the definition of inclusion criterion and -date based on the *second* contact in NPR / RMPS. The inclusion criterion consists of two letters according to which were the first 2 criteria met between prescriptions (detailed as OAD or Insulin) and entry in NPR. Here we make a coarsere clssification based on the *second* criterion met, logically the one met on the day of inclusion:

```
> reg <- transform( reg, inCx = factor(inCr),
+                        inCr = Relevel( factor(inCr),
+                                        list( "DVD"="DVD",
+                                              "Dia"="Dia",
+                                              "Ins"=c("O-I","I-I","N-I"),
+                                              "OAD"=c("O-O","I-O","N-O"),
+                                              "Pod"="Pod",
+                                              "NPR"=c("O-N","I-N","N-N") ) ),
+                         sex = factor( sex, labels=c("M","W") ),
+                        DMtp = factor( DMtp ) )
> reg <- cal.yr( reg )
> cstr( reg )
data.frame 448445 by 22
pnr     1       character
nprtyp  2       character
doNPR   3 cal.yr, numeric
doNPR2  4 cal.yr, numeric
dvdtyp  5       character
doOAD   6 cal.yr, numeric
doIns   7 cal.yr, numeric
doOAD2  8 cal.yr, numeric
doIns2  9 cal.yr, numeric
doPod  10 cal.yr, numeric
doDiaB 11 cal.yr, numeric
sex    12          factor
```

```
doBth  13 cal.yr, numeric
doDth  14 cal.yr, numeric
doDM   15 cal.yr, numeric
doDVD  16 cal.yr, numeric
do2nd  17 cal.yr, numeric
inCr   18          factor
only1  19         numeric
hasdvd 20         numeric
DMtp   21          factor
inCx   22          factor
```

```
> clr <- rainbow(7)[-4][c(1,2,4,5,3,6)]
> save( reg, clr, file = "../nydata/dmreg.rda" )
```

A tabulation shows how the detailed state-inclusion in `inCx` is transformed (grouped) to `inCr`, and how the original classification based on 1st NPR / OAD / Ins relates to that:

```
> system.time( load( file = "../nydata/dmreg.rda" ) )
   user  system elapsed
   1.17    0.07    1.24
```

```
> fCtable( addmargins( with( reg, table( inCx, inCr ) ) ), w=7 )
```

```
     inCr     DVD     Dia     Ins     OAD     Pod     NPR     Sum
inCx
Dia             .  11,468       .       .       .       .  11,468
DVD         5,397       .       .       .       .       .   5,397
I-I             .       .   8,247       .       .       .   8,247
I-N             .       .       .       .       .   1,020   1,020
I-O             .       .       .     450       .       .     450
N-I             .       .  10,176       .       .       .  10,176
N-N             .       .       .       .       .  62,313  62,313
N-O             .       .       .  23,702       .       .  23,702
O-I             .       .   1,237       .       .       .   1,237
O-N             .       .       .       .       .  11,850  11,850
O-O             .       .       . 231,353       .       . 231,353
Pod             .       .       .       .  81,232       .  81,232
Sum         5,397  11,468  19,660 255,505  81,232  75,183 448,445
```

We see that for the `OAD` and `NPR` criteria by far the most common inclusion sequence is two occurrences of the same. Not so for `Ins` where the most common sequence is NPR-Insulin.

3.2 Overview of the register

Here are details on sex, type of diabetes and prevalent at the inception of the register.

```
> # Note the rounding dur to the cal.yr class
> cal.yr(c("1995-12-31","1996-01-01"))
[1] 1995.996 1995.999
attr(,"class")
[1] "cal.yr"  "numeric"
```

```
> fCtable( addmargins( tt <- with( reg, table('>1995'=doDM>1995.998, DMtp, sex ) ),
+                      1:3 ),
+          row.vars=1:2 )
```

```
           sex         M         W       Sum
>1995 DMtp
FALSE T1           12,378     9,596    21,974
      T2           30,338    31,129    61,467
      Sum          42,716    40,725    83,441
TRUE  T1           11,646     8,173    19,819
      T2          192,418   152,767   345,185
      Sum         204,064   160,940   365,004
Sum   T1           24,024    17,769    41,793
      T2          222,756   183,896   406,652
      Sum         246,780   201,665   448,445
```

```
> fCtable( tt, row.vars=1:2 )
```

```
           sex         M         W
>1995 DMtp
FALSE T1           12,378     9,596
      T2           30,338    31,129
TRUE  T1           11,646     8,173
      T2          192,418   152,767
```

We now compute arrays holding the median and IQR for different variables for table 1, classified by diabetes type and sex:

```
> # date of diagnosis
> dd <- with( reg, tapply( doDM,
+                          list( '>1995'=doDM>1995.998, DMtp, sex ),
+                          quantile, 1:3/4 ) )
> # age at diagnosis
> ad <- with( reg, tapply( doDM-doBth,
+                          list( '>1995'=doDM>1995.998, DMtp, sex ),
+                          quantile, 1:3/4 ) )
> # FU time
> fu <- with( reg, tapply( pmin(doDth,2017,na.rm=TRUE)-pmax(doDM,1996),
+                          list( '>1995'=doDM>1995.998, DMtp, sex ),
+                          quantile, 1:3/4 ) )
> # mean date of FU
> df <- with( reg, tapply( (pmin(doDth,2017,na.rm=TRUE)+pmax(doDM,1996))/2,
+                          list( '>1995'=doDM>1995.998, DMtp, sex ),
+                          quantile, 1:3/4 ) )
> # mean age at FU
> af <- with( reg, tapply( (pmin(doDth,2017,na.rm=TRUE)+pmax(doDM,1996))/2-doBth,
+                          list( '>1995'=doDM>1995.998, DMtp, sex ),
+                          quantile, 1:3/4 ) )
> qq <- ZArray( c( list( c("N","Pdx","Adx","Tfu","Pfu","Afu"),
+                        1:3/4 ),
+                  dimnames(ad) ) )
> names( dimnames(qq) )[3]  <- ""
>        dimnames(qq) [[3]] <- c("<1996","1996+")
> str( qq )
 num [1:6, 1:3, 1:2, 1:2, 1:2] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 5
  ..$ : chr [1:6] "N" "Pdx" "Adx" "Tfu" ...
  ..$ : chr [1:3] "0.25" "0.5" "0.75"
  ..$ : chr [1:2] "<1996" "1996+"
  ..$ : chr [1:2] "T1" "T2"
  ..$ : chr [1:2] "M" "W"
```

```
> qq["N",2,,,] <- tt
> qq["Pdx",,,,] <- unlist( dd ) ; qq["Pdx",,'<1996',,] <- 0
> qq["Adx",,,,] <- unlist( ad ) ; qq["Adx",,'<1996',,] <- 0
> qq["Tfu",,,,] <- unlist( fu )
> qq["Pfu",,,,] <- unlist( df )
> qq["Afu",,,,] <- unlist( af )
>       fCtable( qq[ 1,  2     ,,,,drop=FALSE],
+                                    row.vars=c(1,3,5), col.vars=c(4,2), d=0 )
                 T1        T2
                0.5       0.5

N <1996 M     12,378    30,338
        W      9,596    31,129
  1996+ M     11,646   192,418
        W      8,173   152,767
> round( ftable( qq[-1,c(2,1,3),,,], row.vars=c(1,3,5), col.vars=c(4,2) ), 1 )
                 T1                   T2
                0.5   0.25   0.75    0.5   0.25   0.75

Pdx <1996 M     0.0    0.0    0.0    0.0    0.0    0.0
          W     0.0    0.0    0.0    0.0    0.0    0.0
    1996+ M  2005.7 2000.5 2010.9 2008.5 2003.1 2012.4
          W  2005.6 2000.4 2011.0 2008.3 2002.8 2012.3
Adx <1996 M     0.0    0.0    0.0    0.0    0.0    0.0
          W     0.0    0.0    0.0    0.0    0.0    0.0
    1996+ M    30.7   15.7   47.6   62.1   53.0   70.6
          W    26.8   12.4   48.5   64.6   54.1   74.2
Tfu <1996 M    21.0    9.5   21.0    8.9    3.6   18.2
          W    21.0    9.3   21.0    9.1    3.8   18.9
    1996+ M     8.9    4.3   14.5    6.0    3.0   10.5
          W     9.3    4.4   14.9    6.1    3.1   10.9
Pfu <1996 M  2006.5 2000.7 2006.5 2000.4 1997.8 2005.1
          W  2006.5 2000.6 2006.5 2000.6 1997.9 2005.5
    1996+ M  2011.0 2008.0 2013.9 2012.5 2008.8 2014.6
          W  2011.0 2008.1 2014.0 2012.3 2008.7 2014.6
Afu <1996 M    53.5   41.3   65.3   70.9   62.5   78.2
          W    55.5   41.3   71.0   76.0   66.6   82.7
    1996+ M    36.5   20.6   52.4   65.9   57.1   73.9
          W    32.4   17.4   53.8   68.4   58.2   77.7
> fCtable( qq[-1,c(2,1,3),,,], row.vars=c(1,3,5), col.vars=c(4,2), w=7, d=1 )
                  T1                     T2
                 0.5    0.25    0.75     0.5    0.25    0.75

Pdx <1996 M        .       .       .       .       .       .
          W        .       .       .       .       .       .
    1996+ M  2,005.7 2,000.5 2,010.9 2,008.5 2,003.1 2,012.4
          W  2,005.6 2,000.4 2,011.0 2,008.3 2,002.8 2,012.3
Adx <1996 M        .       .       .       .       .       .
          W        .       .       .       .       .       .
    1996+ M     30.7    15.7    47.6    62.1    53.0    70.6
          W     26.8    12.4    48.5    64.6    54.1    74.2
Tfu <1996 M     21.0     9.5    21.0     8.9     3.6    18.2
          W     21.0     9.3    21.0     9.1     3.8    18.9
    1996+ M      8.9     4.3    14.5     6.0     3.0    10.5
          W      9.3     4.4    14.9     6.1     3.1    10.9
Pfu <1996 M  2,006.5 2,000.7 2,006.5 2,000.4 1,997.8 2,005.1
```

```
          W  2,006.5 2,000.6 2,006.5 2,000.6 1,997.9 2,005.5
    1996+ M  2,011.0 2,008.0 2,013.9 2,012.5 2,008.8 2,014.6
          W  2,011.0 2,008.1 2,014.0 2,012.3 2,008.7 2,014.6
Afu <1996 M     53.5    41.3    65.3    70.9    62.5    78.2
          W     55.5    41.3    71.0    76.0    66.6    82.7
    1996+ M     36.5    20.6    52.4    65.9    57.1    73.9
          W     32.4    17.4    53.8    68.4    58.2    77.7
```

3.3 Inclusion criteria

3.3.1 Recorded inclusion dates

While we know that inclusion dates before 1996-01-01 are not usable as proxy for date of diabetes diagnosis, it is illustrative to show how inclusion dates vary by time and criterion.

```
> moi <- with( reg, table( inCr, floor(pmax(doDM,1985.05)*12)/12 ) )
> str( moi )
 'table' int [1:6, 1:384] 7 0 0 0 0 7759 0 0 0 0 ...
 - attr(*, "dimnames")=List of 2
  ..$ inCr: chr [1:6] "DVD" "Dia" "Ins" "OAD" ...
  ..$     : chr [1:384] "1985" "1985.08333333333" "1985.16666666667" "1985.25" ...
> perm <- c(6,4,3,2,1,5)
> plx <- function(ym,leg){
+ barplot( moi[perm,]/1000, beside=FALSE, col=clr[perm],
+          border="transparent", space=0, xaxt="n",
+          ylim=c(0,ym) )
+ mtext( "Monthly no. included (1000s)", las=0, side=2, line=2 )
+ axis( side=1, at=seq(0,(2015-1985)*12,60), labels=seq(1985,2015,5) )
+ axis( side=1, at=seq(0,(2017-1985)*12,12), labels=NA, tcl=-0.3 )
+ abline( v=0:(2017-1985)*12, col=gray(0.7) )
+ abline( v=(1996-1985)*12 )
+ if( leg ) {
+ rect( 2*12,ym,6*12,ym*(13/20), col="white", border="transparent" )
+ text( 3*12, seq(ym*0.95,by=-ym/20,length=6), levels(reg$inCr)[rev(perm)],
+                 col=clr[rev(perm)], font=2, adj=0 ) }
+          }
> par( mfrow=c(2,1), mar=c(3,3,1,0), mgp=c(3,1,0), bty="n", las=1 )
> plx(7,TRUE)
> plx(2.5,FALSE)


> par( mar=c(3,3,1,0), mgp=c(3,1,0), bty="n", las=1 )
> plx(7,TRUE)


> set.seed( 1952 )
> # smooting out the date preference for 1 January
> reg$doDMr <- reg$doDM + runif(nrow(reg),-1,1)/120
> yoi <- with( reg, table( inCr, floor(pmax(doDMr,1985)) ) )
> str( yoi )
 'table' int [1:6, 1:33] 10 0 0 0 0 9003 0 0 0 0 ...
 - attr(*, "dimnames")=List of 2
  ..$ inCr: chr [1:6] "DVD" "Dia" "Ins" "OAD" ...
  ..$     : chr [1:33] "1985" "1986" "1987" "1988" ...
```

```
> par( mar=c(3,3,1,0), mgp=c(3,1,0), bty="n", las=1 )
> plx <- function(ym,leg){
+ barplot( yoi[perm,]/1000, beside=FALSE, col=clr[perm],
+           border="transparent", space=0, xaxt="n",
+           ylim=c(0,ym) )
+ mtext( "Annual no. included (1000s)", las=0, side=2, line=2 )
+ axis( side=1, at=seq(0,(2015-1985),5), labels=seq(1985,2015,5) )
+ axis( side=1, at=seq(0,(2017-1977),1), labels=NA, tcl=-0.3 )
+ abline( v=0:(2017-1977), col=gray(0.7) )
+ abline( v=(1996-1985) )
+ rect( 0,ym,6,ym*(13/20), col="white", border="transparent" )
+ if( leg ) text( 1, seq(ym*0.95,by=-ym/20,length=6),
+                    levels(reg$inCr)[rev(perm)],
+                    col=clr[rev(perm)], font=2, adj=0 )
+           }
> plx(40,TRUE)
```

From figures 3.1 and 3.3 we see the inception of the health services register at 1990-01-02 and the prescription register 1995-01-01. It is clear from the distribution of first dates of OAD and Ins that a substantial fraction of these dates refer to persons with diabetes prior to 1995-01-01, but also that this wears off during 1995, and hence that inclusion dates after 1996-01-01 are reasonably reliable.

It is also pretty clear that inclusion dates from the DADD are largely confined to January and July, so analysis of seasonal patterns of date of diagnosis will be meaningless with the relatively large measurement error in date of diagnosis obtained from DADD.

Finally we see the administrative problems in relation to podiatrists, giving sudden jumps in number of patients included between mid-2002 and mid 2005, as well as the later half of 2012.

3.3.2 Inclusion criteria by year

Next, we make a table of inclusion criteria by sex and calendar year:

```
> crtab <- with( reg, addmargins( table( sex,
+                                        P=pmax(1995,floor(doDM)),
+                                        inCr ),
+                                 margin = c(1,3) ) )#[,,rev(c(5,3,4,6,1,2))]
> dimnames( crtab )[[2]][1] <- "<1996"
> dimnames( crtab )[[2]][-1] <- paste( "", dimnames( crtab )[[2]][-1] )
> str( crtab )
 'table' num [1:3, 1:22, 1:7] 37 27 64 6 5 11 7 2 9 5 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex : chr [1:3] "M" "W" "Sum"
  ..$ P   : chr [1:22] "<1996" " 1996" " 1997" " 1998" ...
  ..$ inCr: chr [1:7] "DVD" "Dia" "Ins" "OAD" ...
> fCtable( crtab, row.vars=1:2, w=7 )
          inCr     DVD     Dia     Ins     OAD     Pod     NPR     Sum
sex P
M   <1996           37       .   2,759  14,116  10,379  15,426  42,717
     1996            6       .     403   4,172   1,009   1,250   6,840
     1997            7       .     385   3,832   1,119   1,208   6,551
     1998            5       .     345   4,237   1,235   1,397   7,219
     1999            6       .     364   4,281   1,332   1,347   7,330
     2000           28       .     351   4,281   1,133   1,420   7,213
     2001           29       .     360   4,816     728   1,454   7,387
```

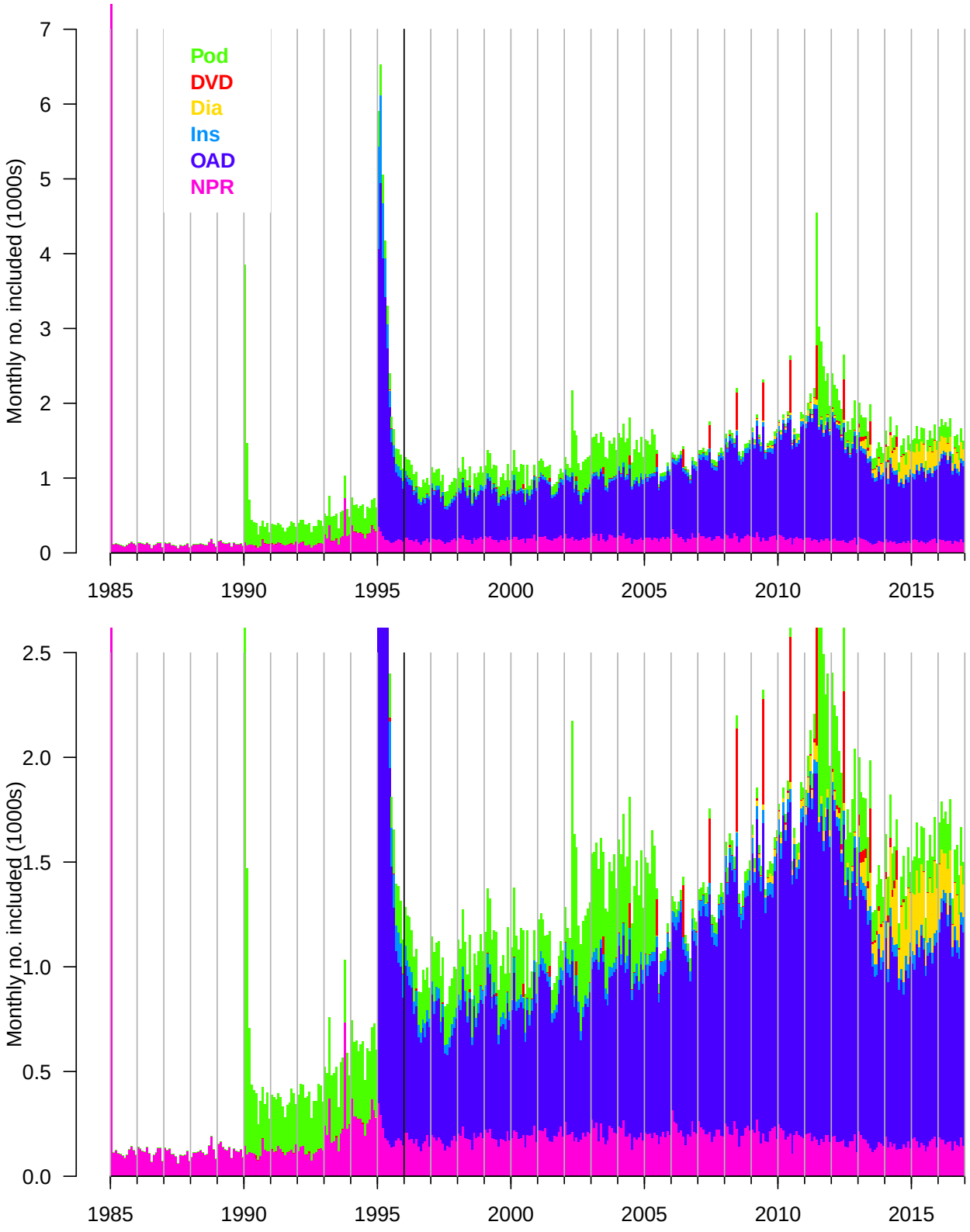


Figure 3.1: *Inclusion in the DMreg by month of inclusion, tick-marks and x-axis labels refer to 1 January each year. The bottom panel is merely a blow-up of the lower quarter of the top one to clarify the patterns. The gray vertical lines indicate 1 January each year, the black vertical line indicates 1996-01-01, wherefrom the register is considered reliable. All dates prior to 1977 are mapped to 1977-01-01 in this plot.* `./graph/reg-hist`

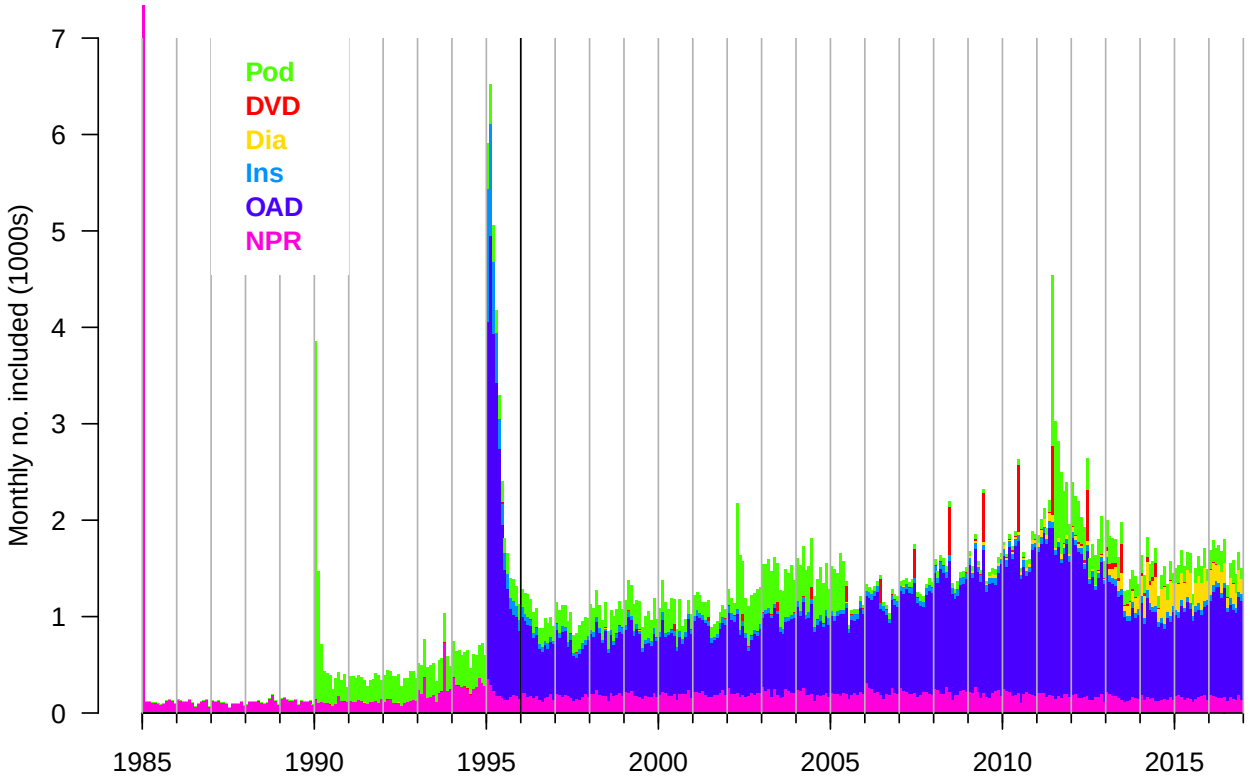


Figure 3.2: *Inclusion in the DMreg by month of inclusion, tick-marks and x-axis labels refer to 1 January each year. The gray vertical lines indicate 1 January each year, the black vertical line indicates 1996-01-01, wherefrom the register is considered reliable. All dates prior to 1985 are mapped to 1985-01-01 in this plot.* `./graph/reg-hist-m`

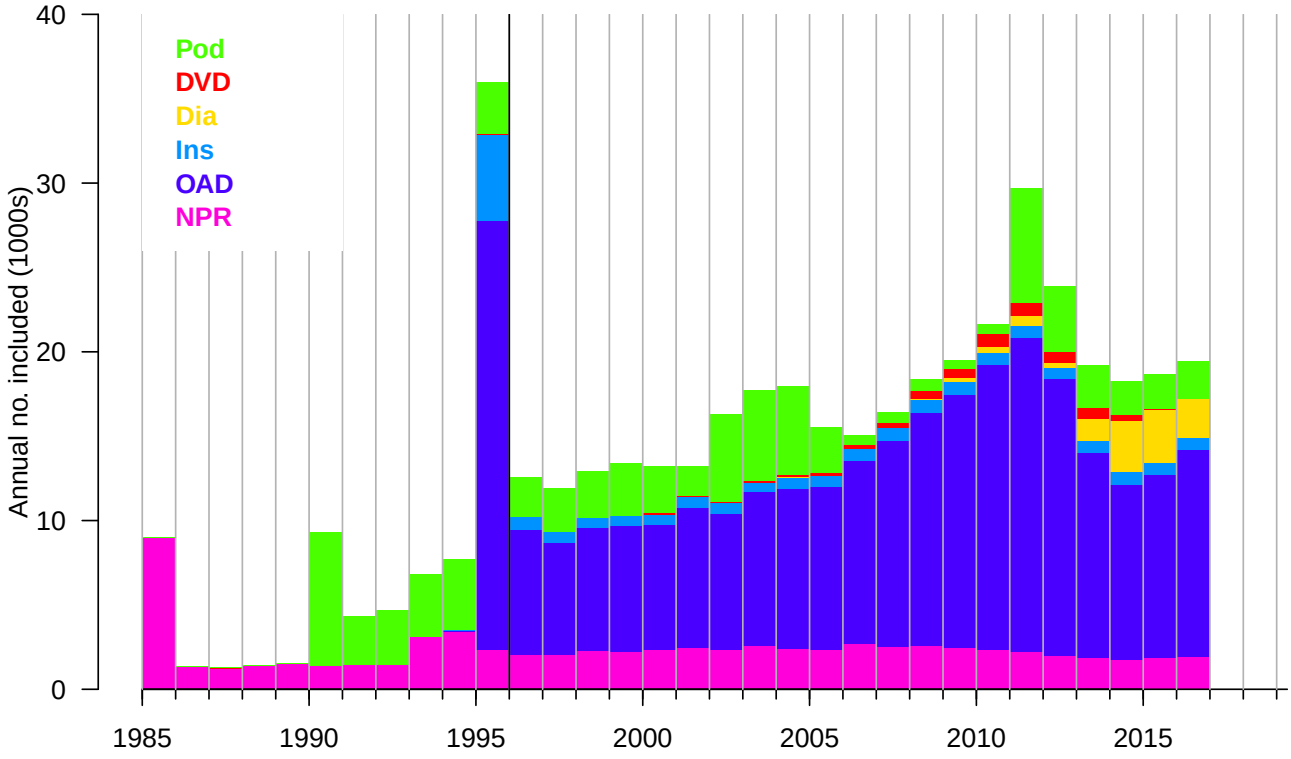


Figure 3.3: *Inclusion in the DMreg by year of inclusion, tick-marks and x-axis labels refer to 1 January each year. The black vertical line indicates 1996-01-01, wherefrom the register is considered reliable. All dates prior to 1977 are mapped to 1977 in this plot.* `./graph/reg-hist-y`

	2002	34	.	360	4,693	2,112	1,469	8,668
	2003	49	.	327	5,399	2,337	1,592	9,704
	2004	69	.	387	5,595	2,293	1,479	9,823
	2005	96	.	373	5,605	1,173	1,444	8,691
	2006	140	.	394	6,367	228	1,625	8,754
	2007	154	.	402	6,920	254	1,549	9,279
	2008	298	.	467	7,872	262	1,573	10,472
	2009	306	158	456	8,706	219	1,552	11,397
	2010	394	195	381	9,723	266	1,499	12,458
	2011	409	328	394	10,643	2,966	1,348	16,088
	2012	307	160	373	9,353	1,893	1,270	13,356
	2013	310	629	438	6,933	1,263	1,165	10,738
	2014	169	1,586	445	6,119	988	1,099	10,406
	2015	2	1,607	395	6,311	1,026	1,182	10,523
	2016	2	1,226	407	7,144	1,155	1,232	11,166
W	<1996	27	.	2,434	11,313	14,596	12,356	40,726
	1996	5	.	371	3,293	1,355	780	5,804
	1997	2	.	271	2,845	1,443	812	5,373
	1998	11	.	250	3,050	1,529	884	5,724
	1999	9	.	250	3,137	1,775	903	6,074
	2000	39	.	265	3,154	1,612	891	5,961
	2001	20	.	284	3,507	1,019	974	5,804
	2002	31	.	292	3,334	3,099	913	7,669
	2003	38	.	246	3,700	3,032	998	8,014
	2004	48	.	287	3,919	2,980	904	8,138
	2005	82	.	295	4,008	1,536	913	6,834
	2006	118	.	307	4,513	329	1,044	6,311
	2007	166	.	334	5,281	398	971	7,150
	2008	220	.	355	5,963	411	1,037	7,986
	2009	227	118	304	6,257	308	923	8,137
	2010	344	167	339	7,116	326	867	9,159
	2011	348	282	293	7,957	3,820	857	13,557
	2012	319	160	266	7,067	1,969	716	10,497
	2013	307	680	314	5,173	1,248	743	8,465
	2014	169	1,455	330	4,217	969	678	7,818
	2015	5	1,587	308	4,479	1,036	736	8,151
	2016	5	1,130	299	5,104	1,072	703	8,313
Sum	<1996	64	.	5,193	25,429	24,975	27,782	83,443
	1996	11	.	774	7,465	2,364	2,030	12,644
	1997	9	.	656	6,677	2,562	2,020	11,924
	1998	16	.	595	7,287	2,764	2,281	12,943
	1999	15	.	614	7,418	3,107	2,250	13,404
	2000	67	.	616	7,435	2,745	2,311	13,174
	2001	49	.	644	8,323	1,747	2,428	13,191
	2002	65	.	652	8,027	5,211	2,382	16,337
	2003	87	.	573	9,099	5,369	2,590	17,718
	2004	117	.	674	9,514	5,273	2,383	17,961
	2005	178	.	668	9,613	2,709	2,357	15,525
	2006	258	.	701	10,880	557	2,669	15,065
	2007	320	.	736	12,201	652	2,520	16,429
	2008	518	.	822	13,835	673	2,610	18,458
	2009	533	276	760	14,963	527	2,475	19,534
	2010	738	362	720	16,839	592	2,366	21,617
	2011	757	610	687	18,600	6,786	2,205	29,645
	2012	626	320	639	16,420	3,862	1,986	23,853
	2013	617	1,309	752	12,106	2,511	1,908	19,203
	2014	338	3,041	775	10,336	1,957	1,777	18,224

```
     2015            7   3,194     703  10,790   2,062   1,918  18,674
     2016            7   2,356     706  12,248   2,227   1,935  19,479
> ( inTp <- dimnames(crtab)[[3]] )
[1] "DVD" "Dia" "Ins" "OAD" "Pod" "NPR" "Sum"
```

Note that the first date category are the persons that are prevalent diabetes cases as of 1 January 1996.

The relative percentages in each year is also of interest:

```
> pctab <- sweep( crtab, 1:2, crtab[,,"Sum"], "/" ) * 100
> fCtable( pctab, row.vars=1:2, w=6, d=1 )
          inCr    DVD    Dia    Ins    OAD    Pod    NPR    Sum
sex P
M   <1996         0.1      .    6.5   33.0   24.3   36.1  100.0
     1996         0.1      .    5.9   61.0   14.8   18.3  100.0
     1997         0.1      .    5.9   58.5   17.1   18.4  100.0
     1998         0.1      .    4.8   58.7   17.1   19.4  100.0
     1999         0.1      .    5.0   58.4   18.2   18.4  100.0
     2000         0.4      .    4.9   59.4   15.7   19.7  100.0
     2001         0.4      .    4.9   65.2    9.9   19.7  100.0
     2002         0.4      .    4.2   54.1   24.4   16.9  100.0
     2003         0.5      .    3.4   55.6   24.1   16.4  100.0
     2004         0.7      .    3.9   57.0   23.3   15.1  100.0
     2005         1.1      .    4.3   64.5   13.5   16.6  100.0
     2006         1.6      .    4.5   72.7    2.6   18.6  100.0
     2007         1.7      .    4.3   74.6    2.7   16.7  100.0
     2008         2.8      .    4.5   75.2    2.5   15.0  100.0
     2009         2.7    1.4    4.0   76.4    1.9   13.6  100.0
     2010         3.2    1.6    3.1   78.0    2.1   12.0  100.0
     2011         2.5    2.0    2.4   66.2   18.4    8.4  100.0
     2012         2.3    1.2    2.8   70.0   14.2    9.5  100.0
     2013         2.9    5.9    4.1   64.6   11.8   10.8  100.0
     2014         1.6   15.2    4.3   58.8    9.5   10.6  100.0
     2015           .   15.3    3.8   60.0    9.8   11.2  100.0
     2016           .   11.0    3.6   64.0   10.3   11.0  100.0
W   <1996         0.1      .    6.0   27.8   35.8   30.3  100.0
     1996         0.1      .    6.4   56.7   23.3   13.4  100.0
     1997           .      .    5.0   52.9   26.9   15.1  100.0
     1998         0.2      .    4.4   53.3   26.7   15.4  100.0
     1999         0.1      .    4.1   51.6   29.2   14.9  100.0
     2000         0.7      .    4.4   52.9   27.0   14.9  100.0
     2001         0.3      .    4.9   60.4   17.6   16.8  100.0
     2002         0.4      .    3.8   43.5   40.4   11.9  100.0
     2003         0.5      .    3.1   46.2   37.8   12.5  100.0
     2004         0.6      .    3.5   48.2   36.6   11.1  100.0
     2005         1.2      .    4.3   58.6   22.5   13.4  100.0
     2006         1.9      .    4.9   71.5    5.2   16.5  100.0
     2007         2.3      .    4.7   73.9    5.6   13.6  100.0
     2008         2.8      .    4.4   74.7    5.1   13.0  100.0
     2009         2.8    1.5    3.7   76.9    3.8   11.3  100.0
     2010         3.8    1.8    3.7   77.7    3.6    9.5  100.0
     2011         2.6    2.1    2.2   58.7   28.2    6.3  100.0
     2012         3.0    1.5    2.5   67.3   18.8    6.8  100.0
     2013         3.6    8.0    3.7   61.1   14.7    8.8  100.0
     2014         2.2   18.6    4.2   53.9   12.4    8.7  100.0
     2015         0.1   19.5    3.8   55.0   12.7    9.0  100.0
```

```
     2016         0.1   13.6    3.6   61.4   12.9    8.5  100.0
Sum <1996         0.1    .      6.2   30.5   29.9   33.3  100.0
     1996         0.1    .      6.1   59.0   18.7   16.1  100.0
     1997         0.1    .      5.5   56.0   21.5   16.9  100.0
     1998         0.1    .      4.6   56.3   21.4   17.6  100.0
     1999         0.1    .      4.6   55.3   23.2   16.8  100.0
     2000         0.5    .      4.7   56.4   20.8   17.5  100.0
     2001         0.4    .      4.9   63.1   13.2   18.4  100.0
     2002         0.4    .      4.0   49.1   31.9   14.6  100.0
     2003         0.5    .      3.2   51.4   30.3   14.6  100.0
     2004         0.7    .      3.8   53.0   29.4   13.3  100.0
     2005         1.1    .      4.3   61.9   17.4   15.2  100.0
     2006         1.7    .      4.7   72.2    3.7   17.7  100.0
     2007         1.9    .      4.5   74.3    4.0   15.3  100.0
     2008         2.8    .      4.5   75.0    3.6   14.1  100.0
     2009         2.7    1.4    3.9   76.6    2.7   12.7  100.0
     2010         3.4    1.7    3.3   77.9    2.7   10.9  100.0
     2011         2.6    2.1    2.3   62.7   22.9    7.4  100.0
     2012         2.6    1.3    2.7   68.8   16.2    8.3  100.0
     2013         3.2    6.8    3.9   63.0   13.1    9.9  100.0
     2014         1.9   16.7    4.3   56.7   10.7    9.8  100.0
     2015         .     17.1    3.8   57.8   11.0   10.3  100.0
     2016         .     12.1    3.6   62.9   11.4    9.9  100.0
```

We can get a graphical overview of the numbers, but first a groom of the labels to get a nice display:

```
> crt <- crtab
> dimnames( crt )[[2]][c(3,4,5,6,
+                        8,9,10,11,
+                        13,14,15,16,
+                        18,19,20,21)] <- ""
> crS <- crt["Sum",,-7]
> crM <- crt["M",,-7]
> crF <- crt["W",,-7]
> names( dimnames(crS) ) <-
+ names( dimnames(crM) ) <-
+ names( dimnames(crF) ) <- NULL
> layout( rbind(c(1,1),2:3), heights=2:1 )
> par( mar=c(1,0,0,0) )
> zz <- mosaicplot( crS, off=0, col=clr, main="", las=1, cex.axis=0.9, border=gray(0.0))
> par( mar=c(1,1,0,0) )
> mosaicplot( crM, off=0, col=clr, main="", las=1, cex.axis=0.9, border=gray(0.0))
> text( 0.95, 0.5, "Men", font=2, col="white", adj=c(1,0), cex=1.5 )
> par( mar=c(1,0,0,1) )
> mosaicplot( crF, off=0, col=clr, main="", las=1, cex.axis=0.9, border=gray(0.0))
> text( 0.95, 0.5, "Women", font=2, col="white", adj=c(1,0), cex=1.5 )
```

The fraction of persons included by each of the criteria varies over time, particularly the fraction included on use of OAD has increased from about half to about 3 quarters, presumably reflecting the increasing intensity of early pharmacological intervention.

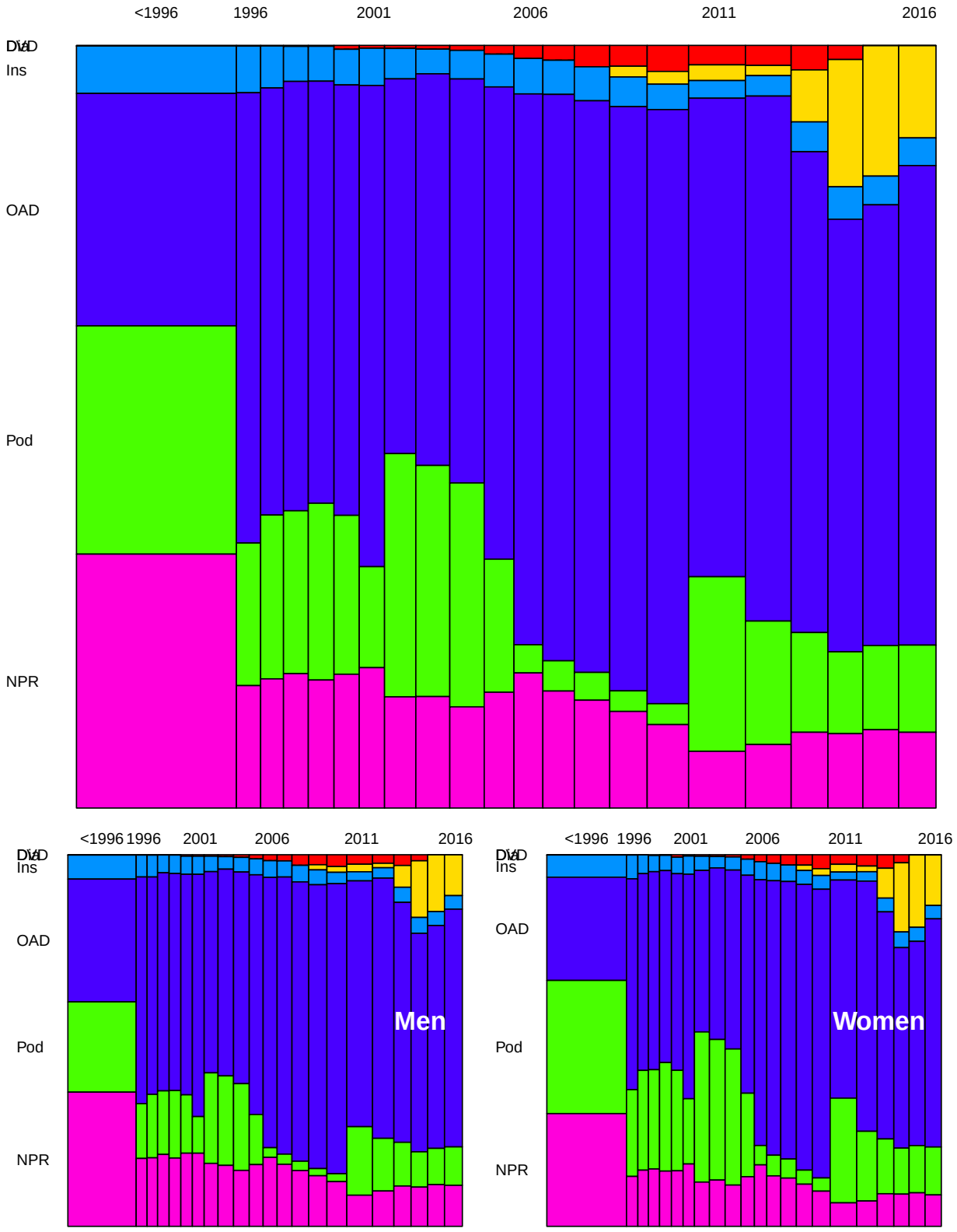


Figure 3.4: *Inclusion criteria by year of inclusion. The widths of the bars are proportional to the number of persons included each year, so the* area *of each rectangle is proportional to the number of persons in it. Colouring of the rectangles is the same as in figure 3.1.* `./graph/reg-mosaic`

3.4 Meeting more than one criterion

It is of interest to see how quickly persons meet the second criterion for inclusion in the register; this will give a hint of possible over-inclusion of persons without diabetes. If there is a very long time to second inclusion for certain persons, this will be an indication of inclusion of more persons that are not diabetes patients.

3.4.1 Defining the follow-up

The analysis will be a survival analysis: How long will persons included in the register "survive" without meeting a 2nd criterion? In the first instance we consider meeting any other than the inclusion criterion as the possible event.

We set up a `Lexis` object to handle this, but first we find the date of meeting the first, resp. 2nd criterion. Here we add a small random quantity (±3 days) to all dates so that no two citeria are met in the same day.

```
> set.seed( 1952 )
> dats <- as.matrix( reg[,c("doIns","doOAD","doDVD","doPod","doDiaB","doNPR")] )
> dats <- dats + runif( dats, -1, 1 )/120
> system.time( idats <- t(apply( dats, 1, sort, na.last=TRUE )) )
   user  system elapsed
  35.08    0.05   35.12
> table( idats[,1] == idats[,2], useNA="ifany" )

 FALSE   <NA>
338474 109971
> apply( idats, 2, function(x) sum( is.na(x) ) )
[1]      0 109971 223143 336172 417018 448445
> reg$doDMr <- idats[,1] # revised date of DM
> reg$do2nd <- idats[,2]
> table( reg$do2nd <= 1996 |
+        reg$doDMr <= reg$doBth |
+        reg$doDMr >= 2017 |
+        reg$doDth <= 1996, useNA="ifany" )

 FALSE   TRUE   <NA>
 82783  54766 310896
```

Nort first that we are not changing the criterion met first, so for those where two criteria are

Note that this definition only is admissible because we only are interested in the probability of meeting the 2nd criterion and not which one it is. In those cases where the second criterion is met close to the first, we will still have the two first dates close to each other.

We have now defined the date of meeting the second criterion, so we can set up a `Lexis` object to hold the follow-up to either 2nd criterion or death.

```
> rL <- Lexis( entry = list( age = pmax(doDMr,1996)-doBth,
+                            per = pmax(doDMr,1996),
+                            tfi = 0 ),
+              exit = list( per = pmin( do2nd, doDth, 2017, na.rm=TRUE ) ),
+       exit.status = factor( 1*(!is.na(do2nd)) +
+                             2*(!is.na(doDth) &
+                                 is.na(do2nd)),
+                             labels=c("One","2+","Dead") ),
+              data = reg )
```

```
NOTE: entry.status has been set to "One" for all.
> nrow( reg ) - nrow( rL )
[1] 54837
> summary( rL )
Transitions:
     To
From    One     2+  Dead  Records:  Events: Risk time:  Persons:
  One 65710 283717 44181    393608   327898    1250762    393608
```

We note that there are some 55,000 persons not in rL, these are the ones that met the second criterion before 1996

```
> summary( rL, by="inCr" )
$DVD

Transitions:
     To
From   One 2+ Dead  Records:  Events: Risk time:  Persons:
  One 4493  0  904      5397      904   36646.61      5397

$Dia

Transitions:
     To
From   One   2+ Dead  Records:  Events: Risk time:  Persons:
  One 4570 6733  165     11468     6898   15286.66     11468

$Ins

Transitions:
     To
From   One    2+ Dead  Records:  Events: Risk time:  Persons:
  One 1687 13296 1700     16683    14996   41717.59     16683

$OAD

Transitions:
     To
From    One     2+  Dead  Records:  Events: Risk time:  Persons:
  One 46404 172468 29122    247994   201590   899787.8    247994

$Pod

Transitions:
     To
From   One    2+  Dead  Records:  Events: Risk time:  Persons:
  One 7500 43679 10996     62175    54675   208580.6     62175

$NPR

Transitions:
     To
From   One    2+ Dead  Records:  Events: Risk time:  Persons:
  One 1056 47541 1294     49891    48835   48742.28     49891
```

```
----------------------
2019-01-06 at 11:56:05
Time elapsed: 00:01:02
----------------------
```

. . . now input from `2nd.tex`

```
--------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-01-06 10:50:40
--------------------------------------
```

3.5 Meeting more than one criterion

It is of interest to see how quickly persons meet the second criterion for inclusion in the register; this will give a hint of possible over-inclusion of persons without diabetes. If there is a very long time to second inclusion for certain persons, this will be an indication of inclusion of more persons that are not diabetes patients.

3.5.1 Defining the follow-up

The analysis will be a survival analysis: How long will persons included in the register "survive" without meeting a 2nd criterion? In the first instance we consider meeting any other than the inclusion criterion as the possible event.

So we load the groomed register:

```
> load( file = "../nydata/dmreg.rda" )
```

We set up a `Lexis` object to handle this, but first we find the date of meeting the first, resp. 2nd criterion. Here we add a small random quantity (±3 days) to all dates so that no two citeria are met in the same day.

```
> set.seed( 1952 )
> dats <- as.matrix( reg[,c("doIns","doOAD","doDVD","doPod","doDiaB","doNPR")] )
> dats <- dats + runif( dats, -1, 1 )/120
> system.time( idats <- t(apply( dats, 1, sort, na.last=TRUE )) )
   user  system elapsed
  40.43    0.12   40.54
> table( idats[,1] == idats[,2], useNA="ifany" )

 FALSE   <NA>
338474 109971
> apply( idats, 2, function(x) sum( is.na(x) ) )
[1]      0 109971 223143 336172 417018 448445
> reg$doDMr <- idats[,1] # revised date of DM
> reg$do2nd <- idats[,2]
> table( reg$do2nd <= 1996 |
+        reg$doDMr <= reg$doBth |
+        reg$doDMr >= 2017 |
+        reg$doDth <= 1996, useNA="ifany" )
```

```
 FALSE   TRUE   <NA>
 82783  54766 310896
```

Note that this definition only is admissible because we only are interested in the probability of meeting the 2nd criterion and not which one it is. In those cases where the second criterion is met close to the first, we will still have the two first dates close to each other.

We have now defined the date of meeting the second criterion, so we can set up a `Lexis` object to hold the follow-up to either 2nd criterion or death.

```
> with( reg, table( D=doDth<=doDMr, '2'=do2nd<=1996, exclude=NULL ) )
       2
D        FALSE   TRUE   <NA>
  FALSE  82783  37145  44188
  TRUE       0      0     72
  <NA>  200934  17612  65711
```

```
> rL <- Lexis( entry = list( age = pmax(doDMr,1996)-doBth,
+                            per = pmax(doDMr,1996),
+                            tfi = 0 ),
+              exit = list( per = pmin( do2nd, doDth, 2017, na.rm=TRUE ) ),
+        exit.status = factor( 1*(!is.na(do2nd)) +
+                              2*(!is.na(doDth) &
+                                 is.na(do2nd)),
+                              labels=c("One","2+","Dead") ),
+              data = subset(reg,doDMr>=1996) )
```

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

```
> nrow( reg ) - nrow( rL )
```

```
[1] 88223
```

```
> summary( rL )
```

```
Transitions:
     To
From    One     2+  Dead  Records:  Events: Risk time:  Persons:
  One 64022 260201 35999    360222   296200    1085104    360222
```

We note that there are some 88,000 persons not in rL, these are the ones that met the second criterion before 1996

```
> zz <- summary( rL, by="inCr" )
> gT <- function( lst ) lst$Transitions
> ll <- do.call( rbind, lapply( zz, gT ) )
> rownames( ll ) <- paste( names(zz), ': ', rownames(ll), sep='' )
> ll
           One     2+  Dead  Records:  Events: Risk time:  Persons:
DVD: One  4448      0   885      5333      885   35369.43      5333
Dia: One  4570   6723   165     11458     6888   15098.95     11458
Ins: One   799  12037  1400     14236    13437   15937.52     14236
OAD: One 46176 156272 24252    226700   180524  807231.62    226700
Pod: One  7095  40244  8399     55738    48643  173845.57     55738
NPR: One   934  44925   898     46757    45823   37620.43     46757
```

There are no transitions to state 2+ from DVD because teh DVD data is only used if no other criterion is met, hence DVD will always be the only criterion met.

3.5.2 Analysis of rate of meeting 2nd criterion

The relevant analysis is of occurrence rates of `2+`, using `tfi` as the primary timescale and age (at inclusion) and date of diagnosis as explanatory variables. The analysis will be conducted separately for each level of `inCr` and sex. However since want the cumulative probability of meeting a second criterion we need to take mortality into account when doing so.

In the next section we also consider the combined endpoint of 2nd criterion and death.

First we split data in first 1 month then 6 month intervals:

```
> system.time( sL <- splitMulti( rL, tfi=c(0:11/12,seq(1,30,1/2)) ) )
   user  system elapsed
  21.86   10.04   27.68
> sL <- transform( sL, ain = age-tfi, pin = per-tfi )
> summary( sL )
Transitions:
     To
From      One     2+  Dead  Records:  Events: Risk time:  Persons:
  One 4540972 260201 35999   4837172   296200    1085104    360222
> range( sL$tfi )
[1]  0.0 20.5
> ( inTp <- c( "All", levels(reg$inCr)[-1] ) )
[1] "All" "Dia" "Ins" "OAD" "Pod" "NPR"
```

We can then fit models for transition rates to `2+` resp. `Dead` separately for the two sexes and also for each inclusion criterion as well as overall:

```
> lmod <- NULL
> system.time(
+ for( sx in c("M","W") )
+ for( dd in c("2+","Dead") )
+ for( ic in inTp )
+    {
+ cat( ic, dd, sx, "running", format(Sys.time(),"%H:%M:%S"), "\n" )
+ lmod <- c( lmod, list(
+   glm( (lex.Xst==dd) ~ Ns( tfi, knots=c(0,0.5,1,2,5,10) ) +
+                        Ns( ain, knots=c(4:8*10) ) +
+                        Ns( pin, knots=seq(1997,2015,,4) ) +
+                        offset( log(lex.dur) ),
+      family = poisson,
+        data = if( ic=="All" ) subset( sL, sex==sx ) else
+                               subset( sL, sex==sx & inCr==ic ) ) ) )
+ names( lmod )[length(lmod)] <- paste( ic, dd, sx, sep="-")
+ gc()
+    } )
> names( lmod )
> save( lmod, file="../nydata/lmod.Rda" )


> system.time( load( file="../nydata/lmod.Rda" ) )
   user  system elapsed
 134.89    4.12  150.42
```

What we want next is to show the cumulative probability of meeting a second criterion separately for men and women and for each *first* entry criterion. This will be done for a person entering the register at age 60 at 1.1.2010 (close to the median age at entry), and additionally to this we will show the RR relative to age 60 and date of inclusion both for 2nd inclusion and death.

So we set up arrays for cumulative risk of a second registration and death without this, as well as the corresponding RRs for the two relative to the reference:

```
> lname <- list( inCr = inTp,
+                Xst = c("2+","Dead"),
+                sex = c("M","W") )
> t.tfi <- seq(0,12,0.1)
> t.ain <- seq(40,80,0.2)
> t.pin <- seq(1996,2017,0.2)
> Rsk <- NArray( c( lname, list( tfi=t.tfi ) ) )
> CHz <-
+ Haz <- NArray( c( lname, list( tfi=t.tfi, cont=c("Est","up","lo") ) ) )
> RRa <- NArray( c( lname, list( age=t.ain, cont=c("Est","up","lo") ) ) )
> RRp <- NArray( c( lname, list( per=t.pin, cont=c("Est","up","lo") ) ) )
> str( CHz )
 logi [1:6, 1:2, 1:2, 1:121, 1:3] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 5
  ..$ inCr: chr [1:6] "All" "Dia" "Ins" "OAD" ...
  ..$ Xst : chr [1:2] "2+" "Dead"
  ..$ sex : chr [1:2] "M" "W"
  ..$ tfi : chr [1:121] "0" "0.1" "0.2" "0.3" ...
  ..$ cont: chr [1:3] "Est" "up" "lo"
```

We need a little extension of the `ci.cum` function in the `Epi` package, allowing a prediction data frame as input to `ci.cum`, so we call it `ci.Cum`[1] and also tailor it to exploit the `Lexis` structure::

```
> ci.Cum <-
+ function( obj, nd, ... )
+   {
+   ci.cum( obj, Epi:::df2ctr(obj,nd),
+           ci.Exp=TRUE, int=mean(nd$lex.dur), ... )[,1:3]
+   }
```

We can now fill in values in these arrays:

```
> for( ic in inTp )
+ for( dd in c("2+","Dead") )
+ for( sx in c("M","W") )
+    {
+ mnam <- paste( ic, dd, sx, sep="-")
+ # cat( mnam, "\n" )
+ Hzd <- data.frame( tfi = t.tfi,
+                    ain = 60,
+                    pin = 2010,
+                lex.dur = 0.1 ) # result is cumulative incidence over interval
+ Haz[ic,dd,sx,,] <- ci.pred( lmod[[mnam]], Hzd )
+ CHz[ic,dd,sx,,] <-  ci.Cum( lmod[[mnam]], Hzd )
+ Ref <- data.frame( tfi = 1,
```

[1] Ultimately to be incorporated in the `Epi` package.

```
+                    ain = 60,
+                    pin = 2010,
+                lex.dur = 1 )
+ aRR <- data.frame( tfi = 1,
+                    ain = t.ain,
+                    pin = 2010,
+                lex.dur = 1 )
+ pRR <- data.frame( tfi = 1,
+                    ain = 60,
+                    pin = t.pin,
+                lex.dur = 1 )
+ RRa[ic,dd,sx,,] <- ci.exp( lmod[[mnam]], list(aRR,Ref) )
+ RRp[ic,dd,sx,,] <- ci.exp( lmod[[mnam]], list(pRR,Ref) )
+    }
```

We have the cumulative hazards over small intervals in `Haz` and the cumulative hazards in `Chz`. From the latter we can construct the so-called "cause-specific risks" of death and 2nd registration with confidence intervals. But we would also like to see the "true risks" taking the competing cause (death) into account:

```
> for( ic in dimnames(Rsk)[[1]] )
+ for( dd in dimnames(Rsk)[[2]] )
+ for( sx in dimnames(Rsk)[[3]] )
+ Rsk[ic,dd,sx,] <- cumsum(Haz[ic,dd,sx,,1]*exp(-CHz[ic,"2+"  ,sx,,1]
+                                                -CHz[ic,"Dead",sx,,1]))
```

We have however not bothered to compute the confidence limits of the latter; the main purpose being to see the discrepancy between the cause-specific and the true probabilities.

We now have the quantities we are interested in; three plots for each combination of sex and event type, death being the secondary one. The plots are the cumulative risk of a second registration for a 60 year old in 2010, and then the cause-specific HRs relative to this.

```
> xclr <- c("black",clr)
> par( mfcol=c(3,2), mar=c(3,1,0.2,0.1), oma=c(0,2,0,0),
+      mgp=c(3,1,0)/1.6, bty="n", las=1 )
> for( sx in c("M","W") )
+    {
+ Rs <- Ch <- NULL
+ for( ic in inTp[-3] ) Ch <- cbind(Ch,CHz[ic,"2+",sx,,])
+ for( ic in inTp[-3] ) Rs <- cbind(Rs,Rsk[ic,"2+",sx, ])
+ matshade( t.tfi, 1-exp(-Ch), col=xclr[-3], lwd=2, plot=TRUE, yaxt="n",
+           ylim=c(0,1), yaxs="i", xlab="Time since inclusion (years)", ylab="" )
+ matlines( t.tfi, Rs, col=xclr[-3], lwd=2, lty="11", lend="butt" )
+ axis( side=2)
+ if( sx=="M") text( 0, seq(0.97,0.70,,6), inTp[-3], font=2, col=xclr[-3], adj=0 )
+ text( 12, 0.03, paste(sx," included at age 60 on 2010/01/01",sep=""), adj=c(1,0) )
+ abline( h=0.5, v=2, col="gray" )
+ axis( side=1, at=0:12, labels=NA, tcl=-0.3 )
+
+ Rs<- NULL
+ for( ic in inTp[-3] ) Rs <- cbind(Rs,RRa[ic,"2+",sx,,])
+ matshade( t.ain, Rs, col=xclr[-3], lwd=2, plot=TRUE,
+           ylim=c(0.2,5), log="y", xlab="Age at inclusion (years)", ylab="" )
+ abline( h=1, v=60, col="gray" )
+ axis( side=1, at=seq(40,80,5), labels=NA, tcl=-0.3 )
+
+ Rs<- NULL
```

```
+ for( ic in inTp[-3] ) Rs <- cbind(Rs,RRp[ic,"2+",sx,,])
+ matshade( t.pin, Rs, col=xclr[-3], lwd=2, plot=TRUE,
+           ylim=c(0.2,5), log="y", xlab="Date of inclusion", ylab="" )
+ abline( h=1, v=2010, col="gray" )
+ axis( side=1, at=1996:2017, labels=NA, tcl=-0.3 )
+    }
> mtext( "Probability of meeting 2nd criterion",
+        side=2, line=1, las=0, outer=TRUE, at=5/6, cex=0.66 )
> mtext( "RR of meeting 2nd criterion",
+        side=2, line=1, las=0, outer=TRUE, at=3/6, cex=0.66 )
> mtext( "RR of meeting 2nd criterion",
+        side=2, line=1, las=0, outer=TRUE, at=1/6, cex=0.66 )
```

From figure 3.5 we see that with the exception of podiatry and insulin, persons included on other criteria as the first have at least 50% probability of meeting a second criterion within two years from meeting the first.

We also see that insulin purchase is the criterion where the difference between the wrongly computed probability and the corretly computed one is largest, simply because the mortality is largest among these.

```
----------------------
2019-01-06 at 10:54:56
Time elapsed: 00:04:16
----------------------
```

. . . now input from `prev.tex`

```
--------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-01-06 11:02:23
--------------------------------------
```

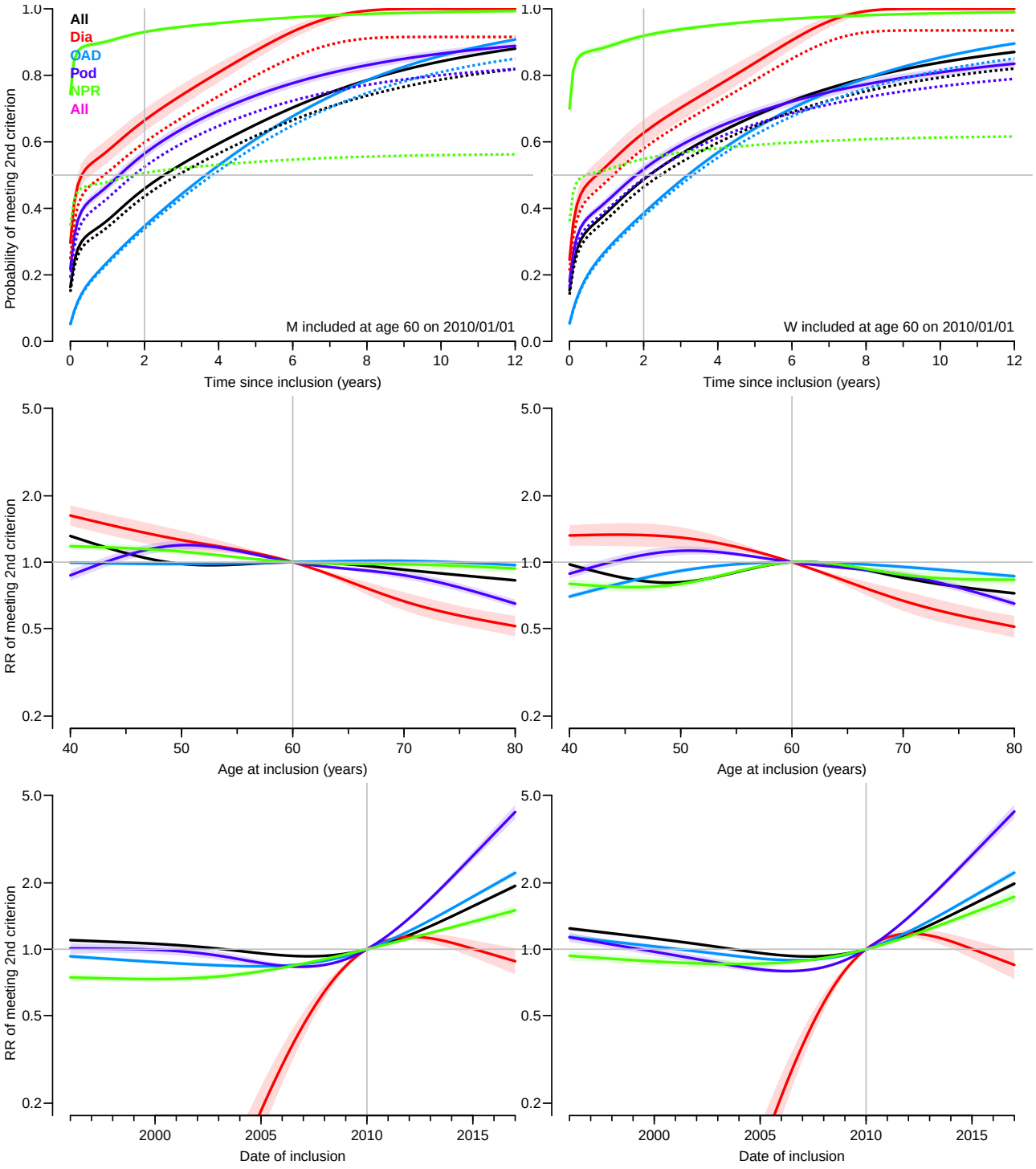


Figure 3.5: *Models for meeting a second criterion after inclusion. This was modeled by a proportional hazards model with effects of time* since *inclusion and age and calendar time* at *inclusion. The full lines are from the model for the occurrence of a second criterion met; in the top panels these represent an over-estimate of the probabilities because death as a competing risk has not been accounted for. The broken lines in the top panels are the probabilities corrected for death, albeit without confidence intervals. The DiaB criterion was so scarce and with limited time that we omitted this from the plot.* `./graph/2nd-Cprob`

Chapter 4

Diabetes prevalence

4.1 Prevalence data

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`:

```
> prN  <- read_sas("../nydata/prv.sas7bdat")
> table( prN$sex )
    1     2
31142 31950
```

```
> head( prN )
```

```
# A tibble: 6 x 7
   pdat   REG state   sex   age     n    a5
  <dbl> <dbl> <chr> <dbl> <dbl> <dbl> <dbl>
1  1996    81 T1        1     2     2     0
2  1996    81 T1        1     3     3     0
3  1996    81 T1        1     4     2     0
4  1996    81 T1        1     6     3     5
5  1996    81 T1        1     7     6     5
6  1996    81 T1        1     8     1     5
```

```
> prN$sex <- factor( prN$sex, labels=c("M","W") )
> fCtable( addmargins( xtabs( n ~ pdat + sex + state, data=prN )[,,-1],
+                      2:3 ),
+          row.vars=1, w=7 )
```

sex	M			W			Sum		
state	T1	T2	Sum	T1	T2	Sum	T1	T2	Sum
pdat									
1996	12,328	30,269	42,597	9,549	31,313	40,862	21,877	61,582	83,459
1997	12,677	33,790	46,467	9,776	34,110	43,886	22,453	67,900	90,353
1998	12,958	36,952	49,910	9,986	36,433	46,419	22,944	73,385	96,329
1999	13,222	40,711	53,933	10,113	39,166	49,279	23,335	79,877	103,212
2000	13,386	44,398	57,784	10,235	42,132	52,367	23,621	86,530	110,151
2001	13,560	47,960	61,520	10,295	44,905	55,200	23,855	92,865	116,720
2002	13,729	51,627	65,356	10,371	47,480	57,851	24,100	99,107	123,207
2003	13,845	56,329	70,174	10,452	51,822	62,274	24,297	108,151	132,448
2004	13,948	61,908	75,856	10,479	56,419	66,898	24,427	118,327	142,754
2005	14,012	67,642	81,654	10,567	61,118	71,685	24,579	128,760	153,339
2006	14,072	72,161	86,233	10,644	64,348	74,992	24,716	136,509	161,225
2007	14,209	76,556	90,765	10,715	66,962	77,677	24,924	143,518	168,442
2008	14,339	81,389	95,728	10,801	70,320	81,121	25,140	151,709	176,849

```
2009        14,485  87,374 101,859  10,901  74,596  85,497  25,386 161,970 187,356
2010        14,648  93,778 108,426  10,979  78,796  89,775  25,627 172,574 198,201
2011        14,745 101,220 115,965  11,078  83,763  94,841  25,823 184,983 210,806
2012        14,860 112,085 126,945  11,177  93,133 104,310  26,037 205,218 231,255
2013        14,988 119,930 134,918  11,289  99,369 110,658  26,277 219,299 245,576
2014        15,116 125,077 140,193  11,458 103,338 114,796  26,574 228,415 254,989
2015        15,304 129,587 144,891  11,614 106,584 118,198  26,918 236,171 263,089
2016        15,512 134,172 149,684  11,826 109,844 121,670  27,338 244,016 271,354
2017        15,684 139,209 154,893  11,930 113,307 125,237  27,614 252,516 280,130
```

4.1.1 Raw tables

Here we produce the prevalence tables needed in the paper:

```
> tt <- xtabs( n ~ pdat + sex + state, data=prN )
> fCtable( mm <- addmargins( tt[,,-1], margin=2:3 ), col.vars=3:2, w=7 )
     state      T1                      T2                     Sum
     sex         M       W     Sum       M       W     Sum       M       W     Sum
pdat
1996        12,328   9,549  21,877  30,269  31,313  61,582  42,597  40,862  83,459
1997        12,677   9,776  22,453  33,790  34,110  67,900  46,467  43,886  90,353
1998        12,958   9,986  22,944  36,952  36,433  73,385  49,910  46,419  96,329
1999        13,222  10,113  23,335  40,711  39,166  79,877  53,933  49,279 103,212
2000        13,386  10,235  23,621  44,398  42,132  86,530  57,784  52,367 110,151
2001        13,560  10,295  23,855  47,960  44,905  92,865  61,520  55,200 116,720
2002        13,729  10,371  24,100  51,627  47,480  99,107  65,356  57,851 123,207
2003        13,845  10,452  24,297  56,329  51,822 108,151  70,174  62,274 132,448
2004        13,948  10,479  24,427  61,908  56,419 118,327  75,856  66,898 142,754
2005        14,012  10,567  24,579  67,642  61,118 128,760  81,654  71,685 153,339
2006        14,072  10,644  24,716  72,161  64,348 136,509  86,233  74,992 161,225
2007        14,209  10,715  24,924  76,556  66,962 143,518  90,765  77,677 168,442
2008        14,339  10,801  25,140  81,389  70,320 151,709  95,728  81,121 176,849
2009        14,485  10,901  25,386  87,374  74,596 161,970 101,859  85,497 187,356
2010        14,648  10,979  25,627  93,778  78,796 172,574 108,426  89,775 198,201
2011        14,745  11,078  25,823 101,220  83,763 184,983 115,965  94,841 210,806
2012        14,860  11,177  26,037 112,085  93,133 205,218 126,945 104,310 231,255
2013        14,988  11,289  26,277 119,930  99,369 219,299 134,918 110,658 245,576
2014        15,116  11,458  26,574 125,077 103,338 228,415 140,193 114,796 254,989
2015        15,304  11,614  26,918 129,587 106,584 236,171 144,891 118,198 263,089
2016        15,512  11,826  27,338 134,172 109,844 244,016 149,684 121,670 271,354
2017        15,684  11,930  27,614 139,209 113,307 252,516 154,893 125,237 280,130
```

```
> str( mm )
 'table' num [1:22, 1:3, 1:3] 12328 12677 12958 13222 13386 ...
 - attr(*, "dimnames")=List of 3
  ..$ pdat : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ sex  : chr [1:3] "M" "W" "Sum"
  ..$ state: chr [1:3] "T1" "T2" "Sum"
```

T1 as percentage of all DM:

```
> fCtable( mm[,,1]/mm[,,3]*100, d=1, w=5 )
     sex     M     W   Sum
pdat
1996      28.9  23.4  26.2
1997      27.3  22.3  24.9
```

```
1998      26.0  21.5  23.8
1999      24.5  20.5  22.6
2000      23.2  19.5  21.4
2001      22.0  18.7  20.4
2002      21.0  17.9  19.6
2003      19.7  16.8  18.3
2004      18.4  15.7  17.1
2005      17.2  14.7  16.0
2006      16.3  14.2  15.3
2007      15.7  13.8  14.8
2008      15.0  13.3  14.2
2009      14.2  12.8  13.5
2010      13.5  12.2  12.9
2011      12.7  11.7  12.2
2012      11.7  10.7  11.3
2013      11.1  10.2  10.7
2014      10.8  10.0  10.4
2015      10.6   9.8  10.2
2016      10.4   9.7  10.1
2017      10.1   9.5   9.9
```

Here is the crude prevalence of diabetes by type and sex:

```
> xx <- addmargins( tt, margin=2:3 )
> fCtable( mm/xx[,,c(4,4,4)]*100, col.vars=c(3:2), d=2, w=5 )
     state    T1                T2                Sum
     sex       M     W   Sum     M     W   Sum     M     W   Sum
pdat
1996        0.47  0.36  0.42  1.16  1.17  1.17  1.64  1.53  1.58
1997        0.48  0.37  0.42  1.29  1.27  1.28  1.78  1.64  1.71
1998        0.49  0.37  0.43  1.41  1.36  1.38  1.90  1.73  1.81
1999        0.50  0.38  0.44  1.54  1.45  1.50  2.05  1.83  1.94
2000        0.51  0.38  0.44  1.68  1.56  1.62  2.19  1.94  2.06
2001        0.51  0.38  0.44  1.81  1.66  1.73  2.32  2.04  2.17
2002        0.52  0.38  0.45  1.94  1.74  1.84  2.45  2.13  2.29
2003        0.52  0.38  0.45  2.11  1.90  2.00  2.62  2.28  2.45
2004        0.52  0.38  0.45  2.31  2.06  2.18  2.83  2.45  2.64
2005        0.52  0.39  0.45  2.52  2.23  2.37  3.04  2.61  2.82
2006        0.52  0.39  0.45  2.68  2.34  2.51  3.20  2.73  2.96
2007        0.52  0.39  0.46  2.83  2.43  2.63  3.35  2.82  3.08
2008        0.53  0.39  0.46  2.99  2.54  2.76  3.52  2.93  3.22
2009        0.53  0.39  0.46  3.19  2.68  2.93  3.72  3.07  3.39
2010        0.53  0.39  0.46  3.41  2.82  3.11  3.94  3.21  3.57
2011        0.53  0.39  0.46  3.66  2.98  3.32  4.19  3.38  3.78
2012        0.54  0.40  0.47  4.04  3.30  3.67  4.57  3.70  4.13
2013        0.54  0.40  0.47  4.30  3.51  3.90  4.84  3.91  4.37
2014        0.54  0.40  0.47  4.46  3.64  4.04  5.00  4.04  4.52
2015        0.54  0.41  0.47  4.59  3.73  4.15  5.13  4.13  4.63
2016        0.54  0.41  0.48  4.70  3.81  4.26  5.24  4.22  4.73
2017        0.54  0.41  0.48  4.83  3.90  4.37  5.38  4.31  4.84
```

4.1.2 Analysis dataset

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

```
> pt1 <- subset( prN, state=="T1"  , select=c(1,2,4,5,6) ) ; names(pt1)[5]<-"T1"
> pt2 <- subset( prN, state=="T2"  , select=c(1,2,4,5,6) ) ; names(pt2)[5]<-"T2"
> pnd <- subset( prN, state=="noDM", select=c(1,2,4,5,6) ) ; names(pnd)[5]<-"nD"
> prN <- merge( pt1, merge( pt2, pnd, all=TRUE ), all=TRUE )
> prN[is.na(prN)] <- 0
> names( prN )[1:4] <- c("P","reg","sex","A")
> prN <- subset( prN, P>1995.5 & A<99.5 )
> prN <- transform( prN, N = T1+T2+nD,
+                        DM = T1+T2 )
> str( prN )
'data.frame':        22000 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","W": 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 2 3 2 0 3 6 1 3 ...
 $ T2 : num  0 0 0 0 0 0 0 0 0 0 ...
 $ nD : num  3822 3939 3930 3872 3685 ...
 $ N  : num  3822 3939 3932 3875 3687 ...
 $ DM : num  0 0 2 3 2 0 3 6 1 3 ...
> fCtable( addmargins(
+          xtabs( cbind(T1,T2,DM,nD,N) ~ P + sex, data=prN ),
+                 margin=2 ),
+          row.vars=2:1, w=11 )
```

sex	P	T1	T2	DM	nD	N
M	1996	12,328	30,266	42,594	2,561,212	2,603,806
	1997	12,677	33,788	46,465	2,569,308	2,615,773
	1998	12,958	36,950	49,908	2,576,079	2,625,987
	1999	13,222	40,709	53,931	2,581,500	2,635,431
	2000	13,386	44,394	57,780	2,586,118	2,643,898
	2001	13,560	47,954	61,514	2,593,059	2,654,573
	2002	13,729	51,621	65,350	2,599,442	2,664,792
	2003	13,845	56,322	70,167	2,603,544	2,673,711
	2004	13,948	61,903	75,851	2,605,319	2,681,170
	2005	14,011	67,635	81,646	2,607,109	2,688,755
	2006	14,072	72,155	86,227	2,611,183	2,697,410
	2007	14,208	76,551	90,759	2,617,327	2,708,086
	2008	14,339	81,384	95,723	2,626,711	2,722,434
	2009	14,485	87,369	101,854	2,638,394	2,740,248
	2010	14,648	93,772	108,420	2,643,350	2,751,770
	2011	14,745	101,215	115,960	2,648,687	2,764,647
	2012	14,860	112,076	126,936	2,649,831	2,776,767
	2013	14,988	119,920	134,908	2,654,598	2,789,506
	2014	15,116	125,068	140,184	2,664,058	2,804,242
	2015	15,304	129,576	144,880	2,680,672	2,825,552
	2016	15,512	134,157	149,669	2,704,151	2,853,820
	2017	15,684	139,192	154,876	2,726,059	2,880,935
W	1996	9,549	31,303	40,852	2,626,243	2,667,095
	1997	9,776	34,100	43,876	2,634,002	2,677,878
	1998	9,986	36,423	46,409	2,640,111	2,686,520
	1999	10,113	39,148	49,261	2,645,846	2,695,107
	2000	10,235	42,118	52,353	2,650,150	2,702,503
	2001	10,295	44,890	55,185	2,656,675	2,711,860
	2002	10,371	47,461	57,832	2,663,147	2,720,979
	2003	10,450	51,802	62,252	2,665,820	2,728,072
	2004	10,477	56,391	66,868	2,667,471	2,734,339

```
     2005       10,566       61,095       71,661    2,669,627    2,741,288
     2006       10,643       64,317       74,960    2,673,817    2,748,777
     2007       10,715       66,924       77,639    2,679,663    2,757,302
     2008       10,801       70,287       81,088    2,687,693    2,768,781
     2009       10,901       74,558       85,459    2,698,575    2,784,034
     2010       10,979       78,757       89,736    2,706,475    2,796,211
     2011       11,076       83,728       94,804    2,713,493    2,808,297
     2012       11,176       93,091      104,267    2,714,711    2,818,978
     2013       11,289       99,319      110,608    2,718,774    2,829,382
     2014       11,458      103,279      114,737    2,726,834    2,841,571
     2015       11,614      106,519      118,133    2,739,417    2,857,550
     2016       11,826      109,769      121,595    2,757,295    2,878,890
     2017       11,930      113,232      125,162    2,776,608    2,901,770
Sum  1996       21,877       61,569       83,446    5,187,455    5,270,901
     1997       22,453       67,888       90,341    5,203,310    5,293,651
     1998       22,944       73,373       96,317    5,216,190    5,312,507
     1999       23,335       79,857      103,192    5,227,346    5,330,538
     2000       23,621       86,512      110,133    5,236,268    5,346,401
     2001       23,855       92,844      116,699    5,249,734    5,366,433
     2002       24,100       99,082      123,182    5,262,589    5,385,771
     2003       24,295      108,124      132,419    5,269,364    5,401,783
     2004       24,425      118,294      142,719    5,272,790    5,415,509
     2005       24,577      128,730      153,307    5,276,736    5,430,043
     2006       24,715      136,472      161,187    5,285,000    5,446,187
     2007       24,923      143,475      168,398    5,296,990    5,465,388
     2008       25,140      151,671      176,811    5,314,404    5,491,215
     2009       25,386      161,927      187,313    5,336,969    5,524,282
     2010       25,627      172,529      198,156    5,349,825    5,547,981
     2011       25,821      184,943      210,764    5,362,180    5,572,944
     2012       26,036      205,167      231,203    5,364,542    5,595,745
     2013       26,277      219,239      245,516    5,373,372    5,618,888
     2014       26,574      228,347      254,921    5,390,892    5,645,813
     2015       26,918      236,095      263,013    5,420,089    5,683,102
     2016       27,338      243,926      271,264    5,461,446    5,732,710
     2017       27,614      252,424      280,038    5,502,667    5,782,705
```

We now have the data in the format we will use for analysis of the prevalences:

```
> save( prN, file="../nydata/prevN.Rda" )
> load(      file="../nydata/prevN.Rda" )
```

4.2 Age-distribution of prevalent diabetes cases

For illustration we show the age-distribution of the diabetes patients:

```
> head( prN )
     P reg sex A T1 T2   nD    N DM
1 1996  81   M 0  0  0 3822 3822  0
2 1996  81   M 1  0  0 3939 3939  0
3 1996  81   M 2  2  0 3930 3932  2
4 1996  81   M 3  3  0 3872 3875  3
5 1996  81   M 4  2  0 3685 3687  2
6 1996  81   M 5  0  0 3784 3784  0
> tt <- addmargins( xtabs( cbind(T1,T2,DM) ~ sex + A,
+                          data = subset( prN, P==2017 & A < 100 ) ),
+                   margin = 1:2 )
> str( tt )
```

```
 'table' num [1:3, 1:101, 1:3] 0 0 0 1 2 3 9 4 13 14 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex: chr [1:3] "M" "W" "Sum"
  ..$ A  : chr [1:101] "0" "1" "2" "3" ...
  ..$    : chr [1:3] "T1" "T2" "DM"
> par( mar=c(3,3,1,0), mgp=c(3,1,0)/1.6, bty='n', las=1 )
> barplot( t(tt["Sum",-101,1:2])/1000, space=0, names.arg=rep("",100),
+          ylim=c(0,10), ylab="No. persons with DM, 1 Jan. 2017 (1000s)",
+          col=cl<-gray(c(2,5)/10), border='transparent', beside=FALSE )
> abline(v=1:9*10, lty="15" )
> axis( side=1, at=1:9*10, col="transparent" )
> text( 27, 10-3:1*0.4, paste( c("T1: ","T2:","All:"), fC(tt["Sum",101,]) ),
+       col=c(cl,gray(0.3)), font=2, adj=1 )
```

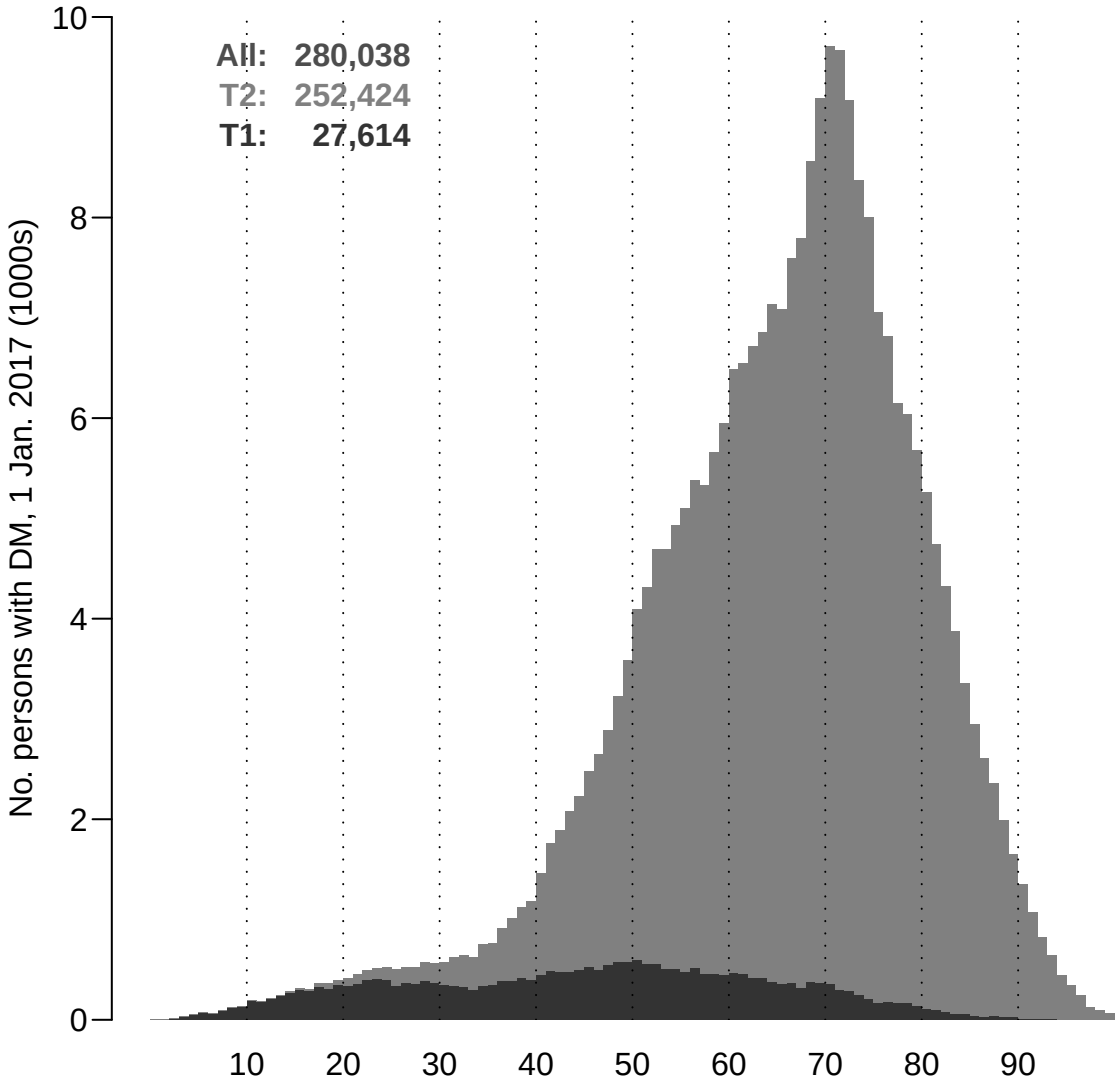


Figure 4.1: *Prevalent number of persons with diabetes in 1-year classes. Dark color is T1, light T2*
./graph/prev-allprev

```
> par( mfrow=c(1,2), mar=c(3,1,1,0), oma=c(0,2,0,0), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> barplot( t(tt["M",-101,1:2])/1000, space=0, names.arg=rep("",100),
+          ylim=c(0,6), ylab="",
+          col=cl<-c("blue","#8888FF"), border='transparent', beside=FALSE )
> text( 27, 6-3:1*0.25, paste( c("T1: ","T2:","  M:"), fC(tt["M",101,]) ),
+       col=c(cl,gray(0.3)), font=2, adj=1 )
> abline(v=1:9*10, lty="15" )
> axis( side=1, at=1:9*10, col="transparent" )
> barplot( t(tt["W",-101,1:2])/1000, space=0, names.arg=rep("",100),
+          ylim=c(0,6), ylab="",
+          col=cl<-c("red","#FF8888"), border='transparent', beside=FALSE )
> text( 27, 6-3:1*0.24, paste( c("T1: ","T2:","  W:"), fC(tt["W",101,]) ),
+       col=c(cl,gray(0.3)), font=2, adj=1 )
> abline(v=1:9*10, lty="15" )
> axis( side=1, at=1:9*10, col="transparent" )
> mtext( "No. persons with DM, 1 Jan. 2017 (1000s)",
+        side=2, line=1, las=0, outer=TRUE )
```

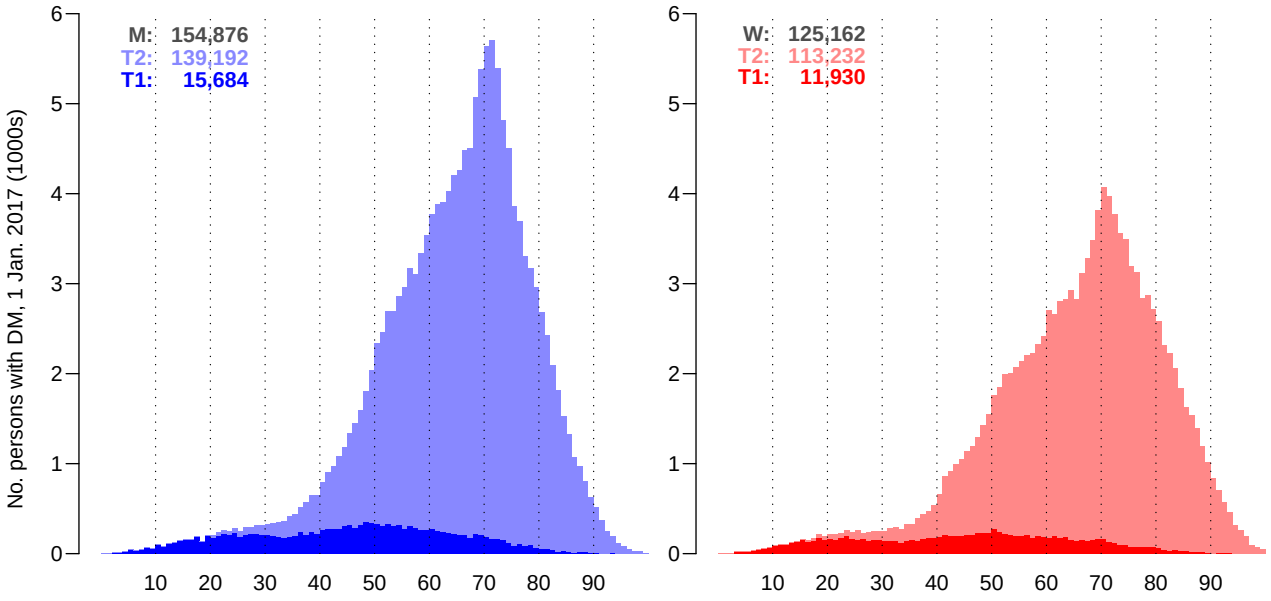


Figure 4.2: *Prevalent number of persons with diabetes at 2017-01-01 in 1-year classes by sex. Men: blue, women: red; dark color is T1, light color T2* `./graph/prev-MFprev`

Note that the peak numbers in figures 4.1 and 4.2 (age 70 at 2017-01-01) correspond to persons born in 1946 — the largest birth cohort ever in Denmark.

4.2.1 Population pyramids

The previous graphs can be put back-to-back to from a population pyramid of obesrved number of cases:

```
> pr.obs <- xtabs( cbind(T1,T2,DM) ~ sex + A + P, data = prN )
> str( pr.obs )
 'xtabs' num [1:2, 1:100, 1:22, 1:3] 0 0 2 3 10 3 8 8 9 8 ...
 - attr(*, "dimnames")=List of 4
```

```
  ..$ sex: chr [1:2] "M" "W"
  ..$ A  : chr [1:100] "0" "1" "2" "3" ...
  ..$ P  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$    : chr [1:3] "T1" "T2" "DM"
 - attr(*, "call")= language xtabs(formula = cbind(T1, T2, DM) ~ sex + A + P, data = prN)
```

With this we can now make a population pyramid for each date 1996-01-01 through 2017-01-01

```
> clr <- c("#7777FF","#0000FF","#FF0000","#FF7777")
> ppyr <-
+ function( pp, lim )
+ {
+ par( mar=c(3,3,3,0), mgp=c(3,1,0)/1.6, las=1 )
+ barplot( height=t( cbind( -pr.obs["M",,pp,"DM"],
+                            pr.obs["M",,pp,"T2"],
+                            pr.obs["M",,pp,"T1"],
+                            pr.obs["W",,pp,"T1"],
+                            pr.obs["W",,pp,"T2"] ) )/ 1000,
+          horiz=TRUE, col=c(NA,clr),
+          border="transparent",space=0,axes=FALSE,
+          names.arg=rep("",dim(pr.obs)[2]),
+          xlim=c(-1,1)*lim*1.05, yaxs="i",
+          xlab="Persons in 1 year class (1000s)",ylab="Age")
+ abline(h=seq(0,100,5),
+        v=seq(-lim,lim,1/2),
+        col=gray(0.9), lty="14", lend="butt" )
+ axis( side=1, at=seq(-lim,lim,2), labels=abs(seq(-lim,lim,2)) )
+ axis( side=1, at=seq(-lim,lim,1), labels=NA, tcl=-0.3 )
+ axis( side=2, at=seq(0,100,20) )
+ axis( side=2, at=seq(0,100,5), labels=NA, tcl=-0.3 )
+ mtext( pp, at=-lim, adj=1.4, cex=1.3, font=1, line=1 )
+ mtext( formatC(sum(pr.obs["M",,pp,"DM"]),0,format="f",big.mark=","),
+        at=-1, col=clr[2], line=0, cex=0.99, adj=1, font=2 )
+ mtext( formatC(sum(pr.obs["W",,pp,"DM"]),0,format="f",big.mark=","),
+        at= 1, col=clr[3], line=0, cex=0.99, adj=0, font=2 )
+ mtext( "N", at=0, line=0, cex=0.99 )
+ }
> pdf( "./graph/prev-obs-film.pdf", width=8, height=6 )
> for( pp in paste(1996:2017) ) ppyr( pp, lim=6 )
> dev.off()
null device
          1
```

4.3 Models for prevalence

We model the prevalences as of 1 January each of the years 1996–2017, 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 knots for T1, T2 and DM:

```
>   kn1 <- c( 10, 20, 40, 60 )
> ( kn2 <- c( 15, with( prN, quantile( rep(A,T2), qn(10) ) ) ) )
     5% 15% 25% 35% 45% 55% 65% 75% 85% 95%
 15  43  52  57  61  65  68  71  75  79  86
```

```
> ( knd <- c( 15, with( prN, quantile( rep(A,DM), qn(10) ) ) ) )
     5% 15% 25% 35% 45% 55% 65% 75% 85% 95%
 15  33  47  54  59  63  66  70  74  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 prevalences (for each 0.5 year) in an array, we set up — note that we also make room for the `gam` fitting:

```
> # predictins at midpoints of months used in components
> m.pt <- seq(0,100,1/12)[-1]-1/24
> pr.ini <-
+ pr.fin <- NArray( list( typ = c("T1","T2","DM"),
+                          sex = c("M","W","B"), #levels( prN$sex ),
+                            A = m.pt,
+                                c("Est","lo","hi") ) )
> A.pt <- seq(0,100,0.5)
> T.pt <- 1996:2017
> parr <- NArray( list( mod = c("glm","gam"),
+                       typ = c("T1","T2","DM"),
+                       sex = c("M","W","B"), #levels( prN$sex ),
+                         A = A.pt,
+                         T = T.pt,
+                             c("Est","lo","hi") ) )
> tarr <- NArray( dimnames(parr)[c(1:3,6)] )
> str( parr )
 logi [1:2, 1:3, 1:3, 1:201, 1:22, 1:3] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 6
  ..$ mod: chr [1:2] "glm" "gam"
  ..$ typ: chr [1:3] "T1" "T2" "DM"
  ..$ sex: chr [1:3] "M" "W" "B"
  ..$ A  : chr [1:201] "0" "0.5" "1" "1.5" ...
  ..$ T  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$    : chr [1:3] "Est" "lo" "hi"
> str( tarr )
 logi [1:2, 1:3, 1:3, 1:3] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 4
  ..$ mod: chr [1:2] "glm" "gam"
  ..$ typ: chr [1:3] "T1" "T2" "DM"
  ..$ sex: chr [1:3] "M" "W" "B"
  ..$    : chr [1:3] "Est" "lo" "hi"
> mod <- list( )
> length(mod) <- 9
> dim     ( mod ) <- dim     ( parr )[2:3]
> dimnames( mod ) <- dimnames( parr )[2:3]
> akn <- mod
> str( mod )
List of 9
 $ : NULL
 $ : NULL
 $ : NULL
 $ : NULL
 $ : NULL
 $ : NULL
 $ : NULL
 $ : NULL
```

```
 $ : NULL
 - attr(*, "dim")= Named int [1:2] 3 3
  ..- attr(*, "names")= chr [1:2] "typ" "sex"
 - attr(*, "dimnames")=List of 2
  ..$ typ: chr [1:3] "T1" "T2" "DM"
  ..$ sex: chr [1:3] "M" "W" "B"
```

The `tarr` array is for storing the overall time trend in prevalence, as estimated in a model with log-linear trend.

The following loop fits models and fill estimates and predictions in them:

```
> system.time(
+ for( tp in dimnames(parr)[["typ"]] )
+    {
+    prN$X <- prN[,tp]
+    a.kn <- switch( tp, T1 = kn1,
+                        T2 = kn2,
+                        DM = knd )
+ for( sx in dimnames(parr)[["sex"]] )
+    {
+ for( it in T.pt )
+    {
+    # separate models for each year
+    mp <- glm( cbind(X,N-X) ~ Ns(A,knots=a.kn),
+               family = binomial( link=log ),
+                 data = subset( prN, (if(sx!="B") sex==sx else TRUE) & P==it ) )
+    parr["glm",tp,sx,,paste(it),] <- ci.pred( mp, newdata=data.frame(A=A.pt) )
+    if( it == 1996 ) pr.ini[tp,sx,,] <- ci.pred( mp, newdata=data.frame(A=m.pt) )
+    if( it == 2017 )
+      {
+                     pr.fin[tp,sx,,] <- ci.pred( mp, newdata=data.frame(A=m.pt) )
+      mod[tp,sx] <- list(mp)
+      akn[tp,sx] <- list(a.kn)
+      }
+    mp <- gam( cbind(X,N-X) ~ s(A),
+               family = binomial( link=log ),
+                 data = subset( prN,(if(sx!="B") sex==sx else TRUE) & P==it ) )
+    parr["gam",tp,sx,,paste(it),] <- ci.pred( mp, newdata=data.frame(A=A.pt) )
+    }
+    # model for all years
+    mt <- glm( cbind(X,N-X) ~ Ns(A,knots=a.kn) + P,
+               family = binomial( link=log ),
+                 data = subset( prN, (if(sx!="B") sex==sx else TRUE) ) )
+    tarr["glm",tp,sx,] <- ci.exp( mt, subset="P" )
+    mt <- gam( cbind(X,N-X) ~ s(A) + P,
+               family = binomial( link=log ),
+                 data = subset( prN, (if(sx!="B") sex==sx else TRUE) ) )
+    tarr["gam",tp,sx,] <- ci.exp( mt, subset="P" )
+    }
+    } )
   user  system elapsed
 237.91    3.62  241.55
```

We then save the observed (`pr.obs`) and predicted prevalences (`parr`, `pr.ini`, `pr.fin`)

```
> round( range( parr, na.rm=T ), 4 )
[1] 0.0000 0.1933
```

```
> save( parr, mod, akn, pr.obs, pr.ini, pr.fin, file='../nydata/prevalences.Rda' )
```

Here are the average annual changes in prevalence by sex, type and model:

```
> round( ftable( (tarr-1)*100, col.vars=c(1,4) ), 2 )
        mod  glm            gam
             Est   lo   hi  Est   lo   hi
typ sex
T1  M       0.51 0.46 0.57 0.46 0.40 0.51
    W       0.52 0.46 0.59 0.47 0.41 0.54
    B       0.53 0.48 0.57 0.47 0.43 0.51
T2  M       5.64 5.61 5.66 5.64 5.61 5.66
    W       5.22 5.19 5.24 5.22 5.19 5.24
    B       5.47 5.45 5.49 5.47 5.45 5.49
DM  M       4.82 4.80 4.84 4.81 4.79 4.83
    W       4.55 4.53 4.58 4.55 4.52 4.57
    B       4.72 4.71 4.74 4.72 4.70 4.73
```

Thus we see that the overall trend in T1D prevalence is very modest, (about 0.47%/year) as compared to the change in T2 prevalence (about 5.5%/year).

We want to be able to show only every 7$^{\text{th}}$ date of prevalence so we need:

```
> ( wh <- paste(seq(1996,2017,7)) )
[1] "1996" "2003" "2010" "2017"
```

The prevalences of all diabetes for each of the 22 dates 1 Jan 1996–2017 are shown in figures 4.3 and 4.4:

```
> plpr <- function( mod ){
+ par( mfrow=c(1,2), mar=c(3,0,0,0), oma=c(0,3,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ for( sx in dimnames(parr)[["sex"]][1:2] )
+    {
+    matplot( A.pt, parr[mod,"DM",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 20),
+             yaxs="i", xaxt="n", yaxt="n" )
+    axis( side=1, at=seq(10,90,20) )
+    axis( side=1, at=1:9*10, tcl=-0.4, labels=NA )
+    axis( side=1, at=1:19*5, tcl=-0.3, labels=NA )
+    if( sx=="M" ) { axis( side=2 )
+    axis( side=2, at=0:20, labels=NA, tcl=-0.3 ) }
+    abline( v=80, col=gray(0.9) )
+    mtext( "Age", side=1, line=3/1.6 )
+    }
+ mtext( "Prevalence of diabetes (%)", side=2, line=3/1.6, outer=TRUE, las=0 )
+ }
> plpr( "glm" )


> plpr( "gam" )
```

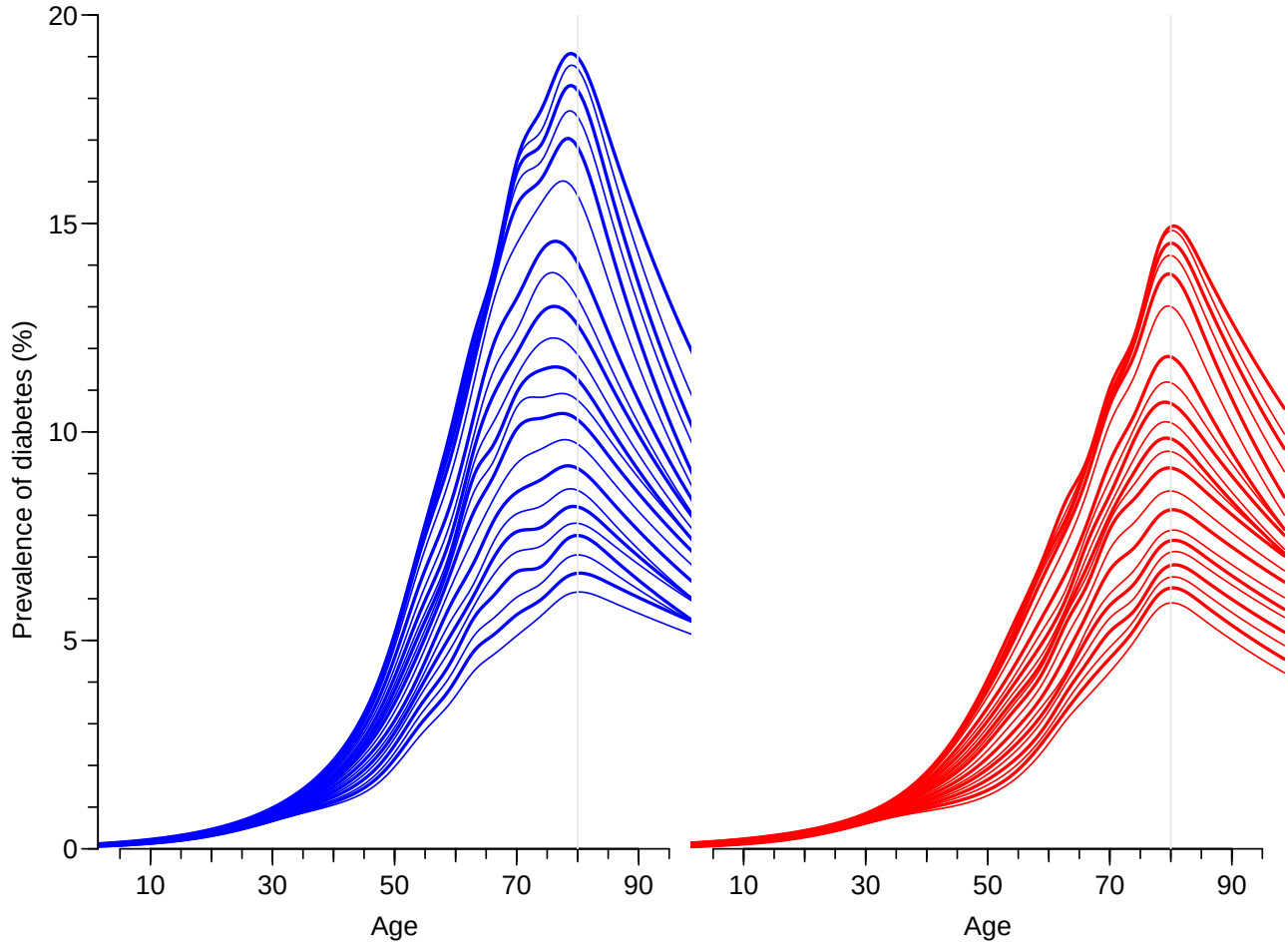


Figure 4.3: *Age-specific prevalences for each of the dates 1996-01-01, 1997-01-01, . . . 2017-01-01 for T1 and T2 combined, using* `glm` *models with explicit knots. Blue curves are men, red women. Thin curves are even years, thick curves are odd years.* `./graph/prev-dm2sx-glm`

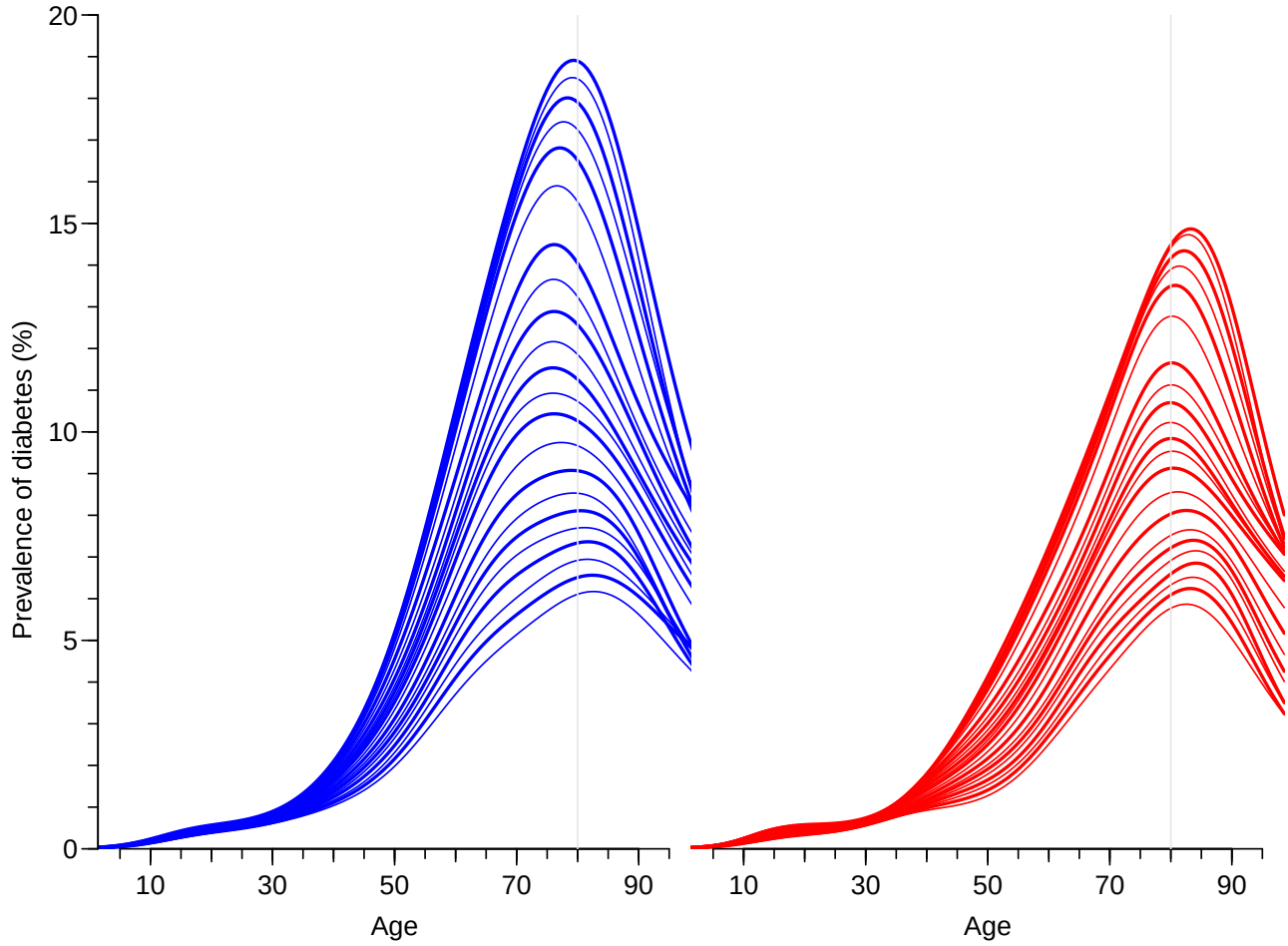


Figure 4.4: *Age-specific prevalences for each of the dates 1996-01-01, 1997-01-01, . . . 2017-01-01 for T1 and T2 combined, using* `gam` *models. Blue curves are men, red women. Thin curves are even years, thick curves are odd years.* `./graph/prev-dm2sx-gam`

The prevalences subdivided by type of diabetes are shown for the same 22 dates in figures 4.5 and 4.6.

```
> plpr <- function(mod){
+ 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"]][1:2] )
+    {
+    matplot( A.pt, parr[mod,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 20),
+             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)
+      axis( side=2, at=if( tp=="T1" ) 1:9/10 else 1:19, labels=NA, tcl=-0.3 )
+      mtext( paste( tp, "diabetes prevalence (%)"), side=2, line=2, las=0 )
+      }
+    }
+ 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 )
+ }
> plpr( "glm" )


> plpr( "gam" )
```

Simplified versions with only 4 dates, but with confidence limits are shown in figures 4.7 and 4.8.

```
> p22 <- function(wh,mod){
+ 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" )
+ whl <- 0
+ for( tp in dimnames(parr)[["typ"]][1:2] )
+ for( sx in dimnames(parr)[["sex"]][1:2] )
+    {
+    whl <- whl + 1
+    matplot( A.pt, parr[mod,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 20),
+             yaxs="i", xaxt="n", yaxt="n" )
+    text( 5, if(tp=="T1") 0.95 else 19.5, letters[whl], font=2, adj=0, cex=1.3)
+    for( i in wh )
+       polygon( c(A.pt,rev(A.pt)), c(parr[mod,tp,sx,,i,2],
+                                     rev(parr[mod,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(61,length(wh)), txpos <- seq(0.55,0.25,,length(wh))-0.2*(sx=="W"), wh, adj=0 )
+    segments( rep(40,5), parr[mod,tp,sx,"40",wh,1]*100,
+              rep(60,5), txpos )
+    segments( rep(85,5), parr[mod,tp,sx,"85",wh,1]*100,
+              rep(70,5), txpos ) } else {
```

```
+    text( rep(90,5), parr[mod,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 )
+      mtext( paste( tp, "diabetes prevalence (%)"), side=2, line=2, las=0 )
+      }
+    }
+ 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( paste(seq(2017,1996,-7)), "glm" )


> p22( paste(seq(2017,1996,-7)), "gam" )
```

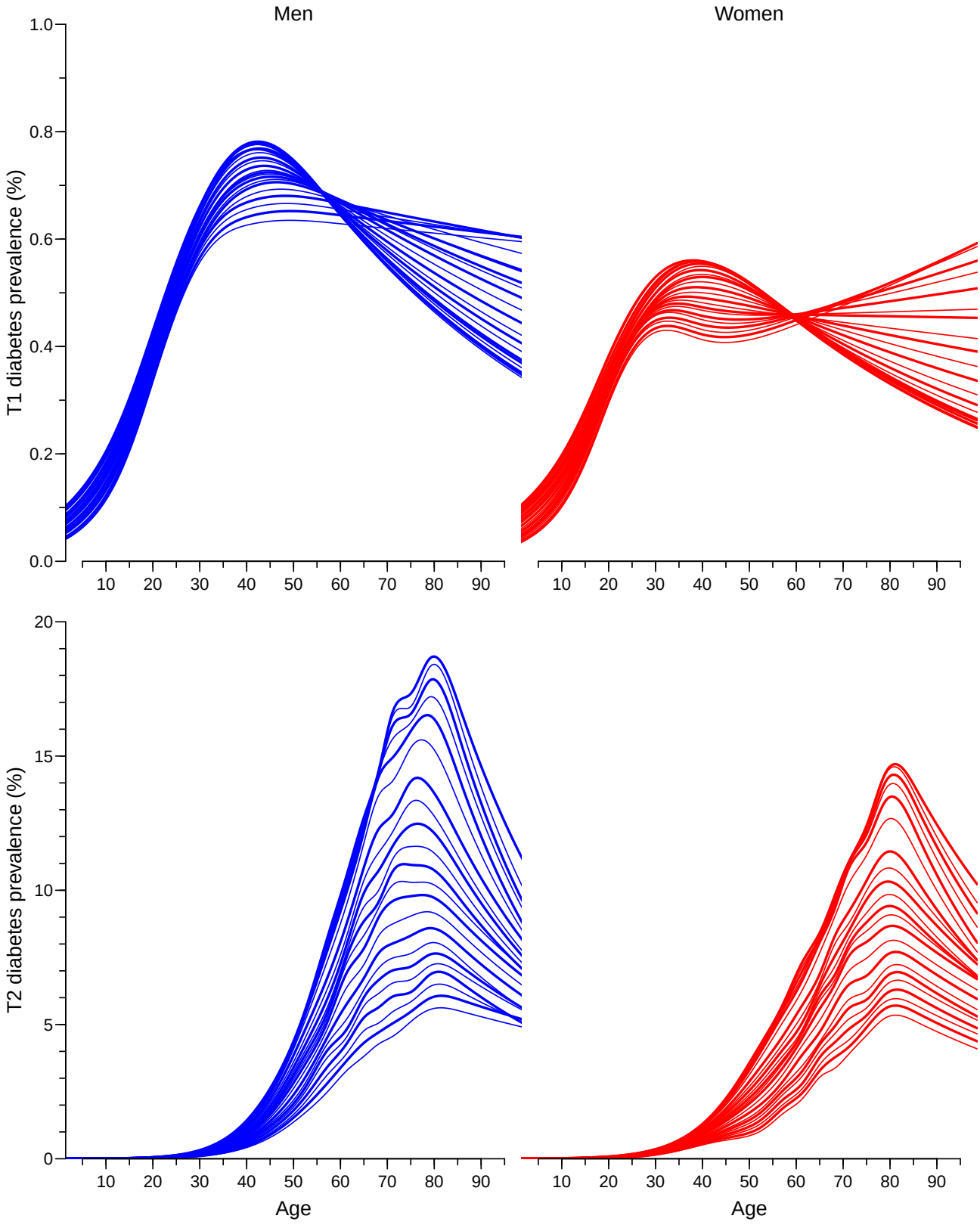


Figure 4.5: *Age-specific prevalences for each of the dates 1996-01-01, 1997-01-01, . . . 2016-01-01 for T1 and T2, using* `glm` *models with 4 (T1) and 11 (T2) (slightly arbitrary) knots for age. Blue curves are men, red women. Thin curves are even years, thick odd years..*`/graph/prev-2tp2sx-glm`

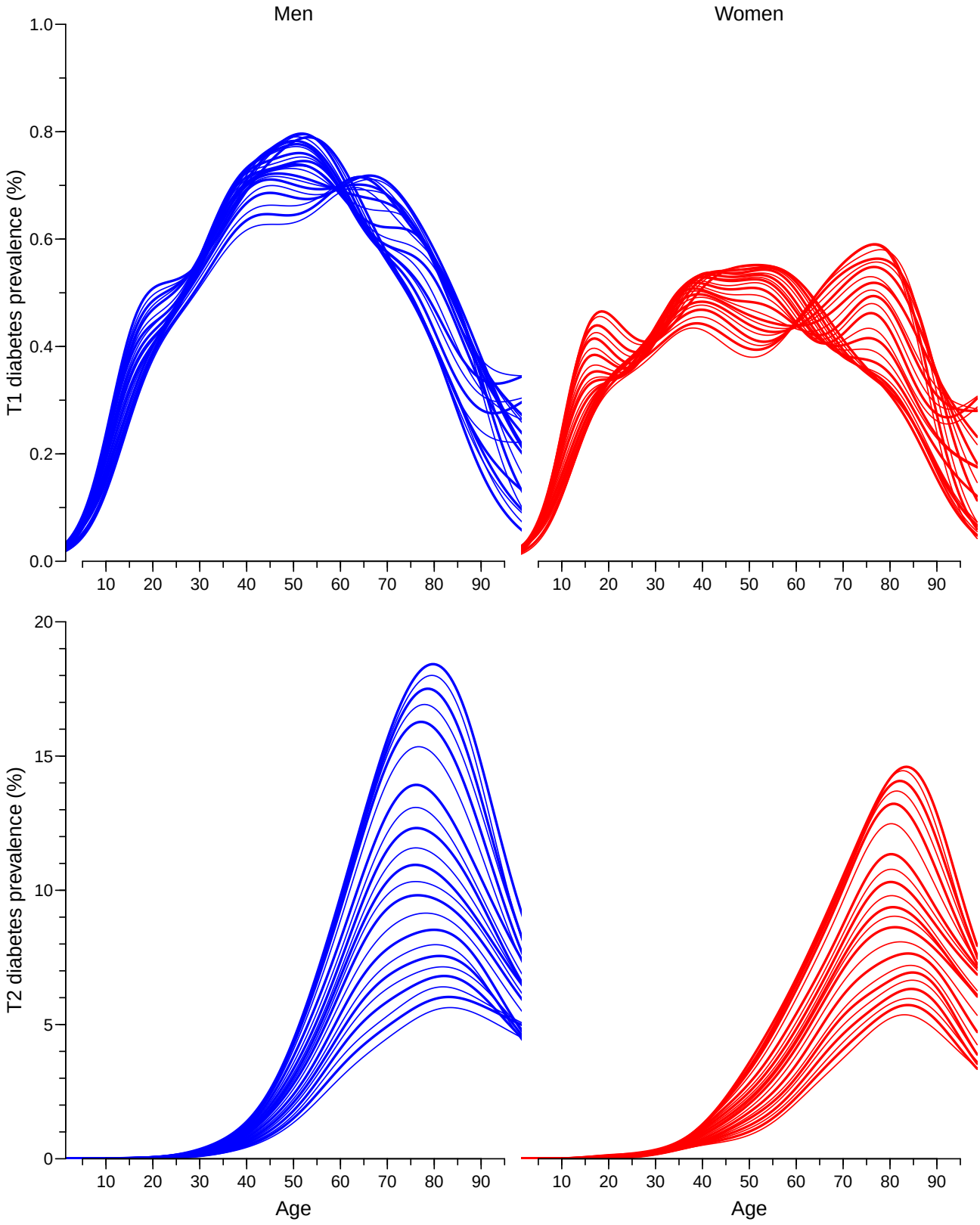


Figure 4.6: *Age-specific prevalences for each of the dates 1996-01-01, 1997-01-01, . . . 2016-01-01 for T1 and T2, using* `gam` *models. Blue curves are men, red women. Thin curves are even years, thick odd years.* `./graph/prev-2tp2sx-gam`

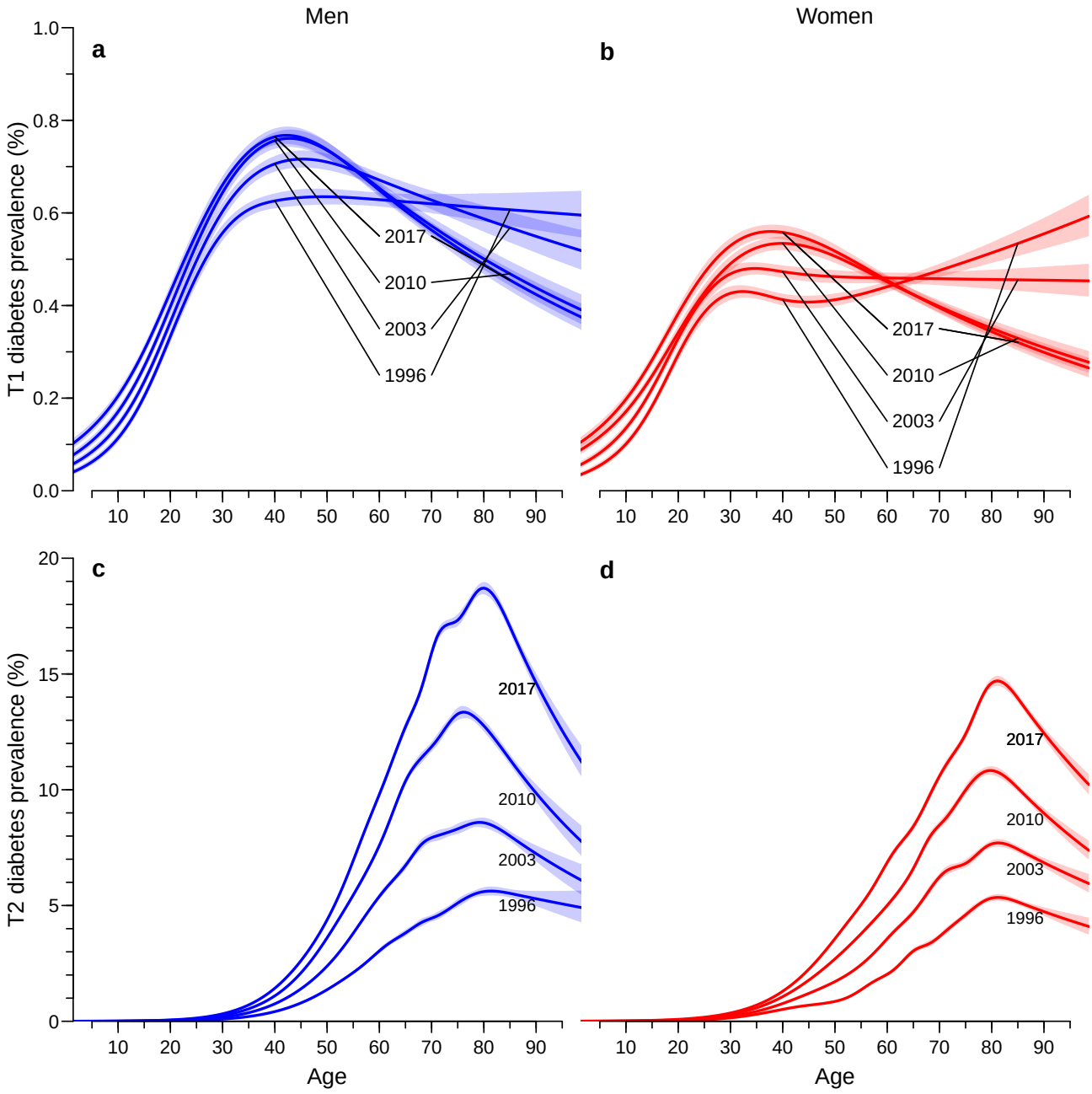


Figure 4.7: *Age-specific prevalences for each of the dates 1997-01-01, 2001-01-01, . . . 2017-01-01 for T1 and T2 diabetes, using* `glm` *models with explicit knots. Blue curves are men, red women, shaded areas are 95% confidence intervals.* `./graph/prev-2tp2sxci-glm`

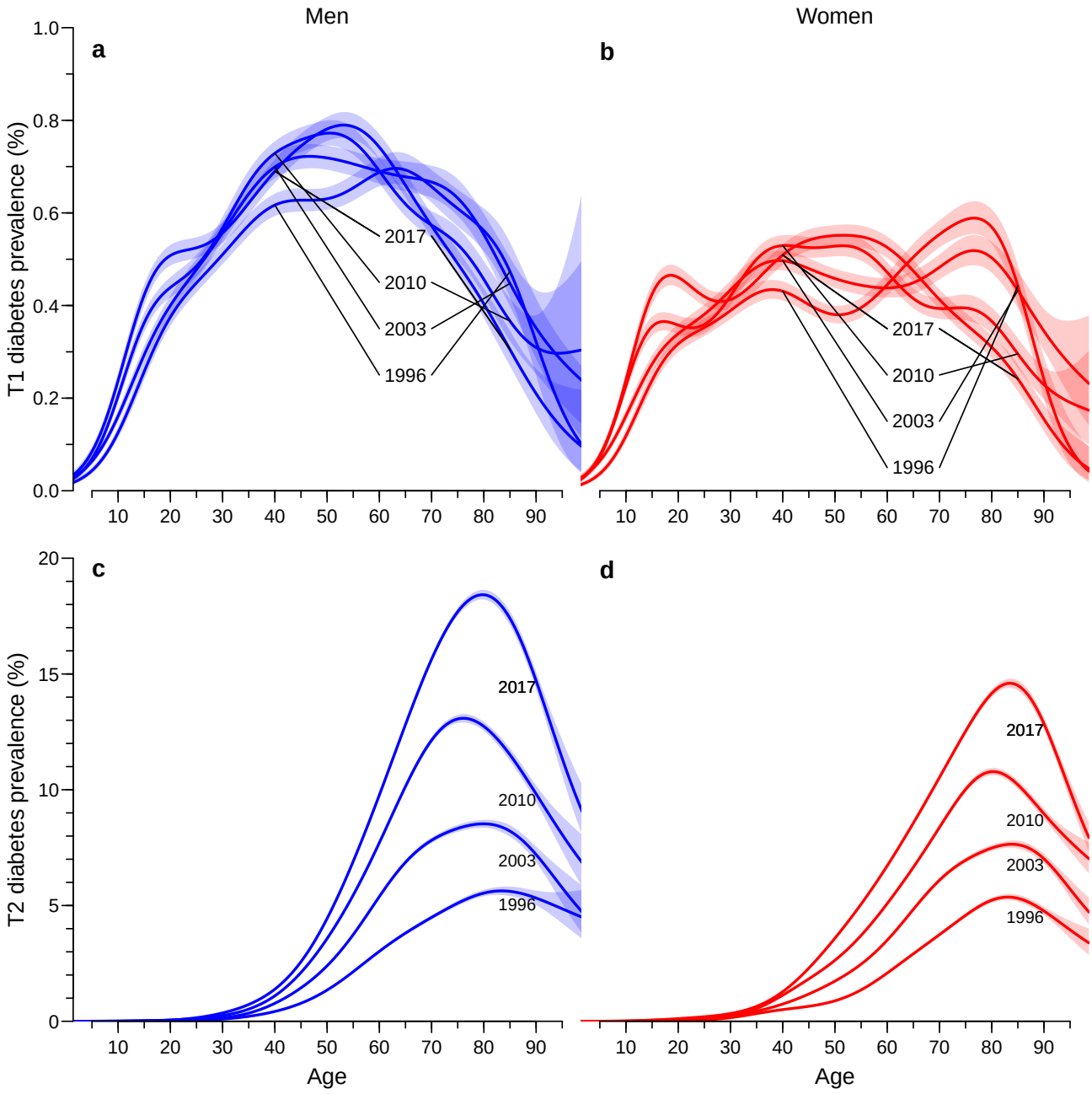


Figure 4.8: *Age-specific prevalences for each of the dates 1997-01-01, 2001-01-01, . . . 2017-01-01 for T1 and T2 diabetes, using* `gam` *models. Blue curves are men, red women, shaded areas are 95% confidence intervals.* `./graph/prev-2tp2sxci-gam`

4.3.1 Fration of patients with T1D

Finally, we present the relative fraction of all diabetes that is T1 as a function of age:

```
> prfr <- function( mod ){
+ par( mfrow=c(1,2), mar=c(3,0,0,0), oma=c(0,3,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ for( sx in dimnames(parr)[["sex"]] )
+    {
+    matplot( A.pt, parr[mod,"T1",sx,,,1]/
+                  (parr[mod,"T1",sx,,,1]+
+                   parr[mod,"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 )
+ }
> prfr( "glm" )
```

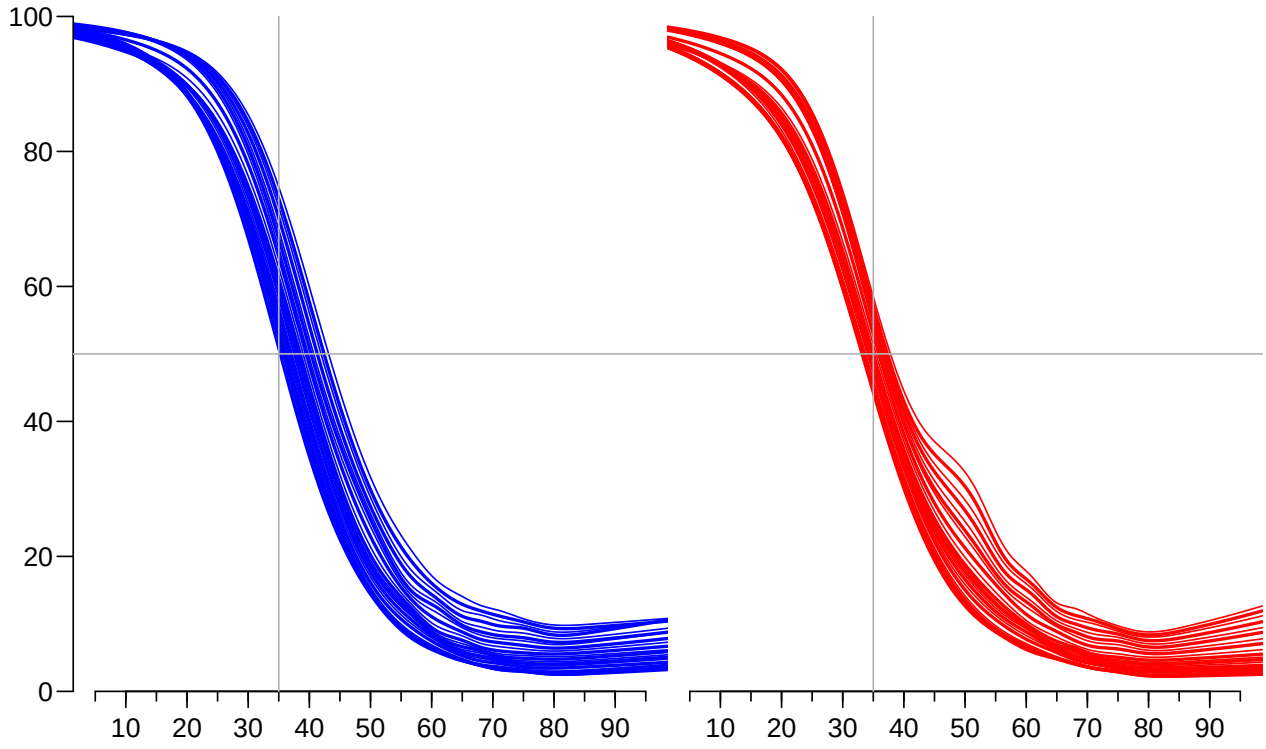


Figure 4.9: *Fraction of T1D among all diabetes patients for dates 1996-01-01, 1997-01-01 etc., based on* `glm` *models.* `./graph/prev-p-ratio-glm`

```
> prfr( "gam" )
```

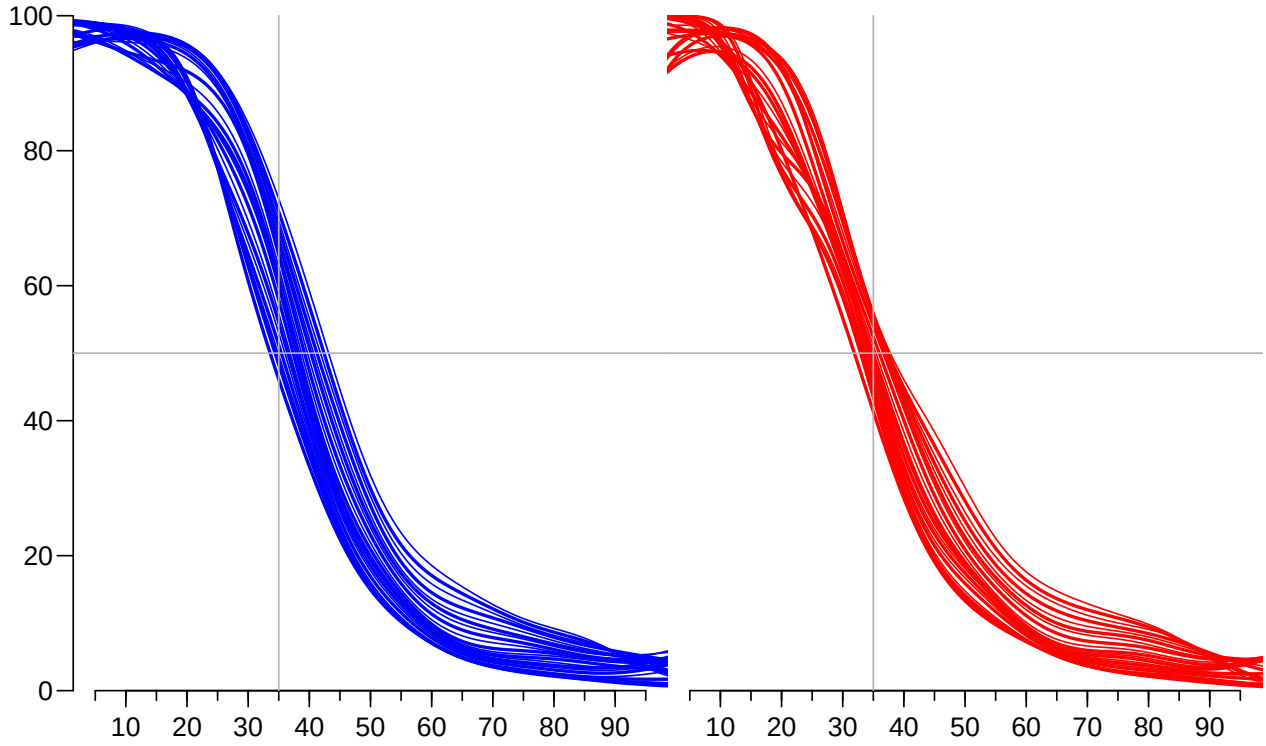


Figure 4.10: *Fraction of T1D among all diabetes patients for dates 1996-01-01, 1997-01-01 etc., based on* `gam` *models.* `./graph/prev-p-ratio-gam`

```
> prfr <- function( mod ){
+ par( mfrow=c(1,2), mar=c(3,0,0,0), oma=c(0,3,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ for( sx in dimnames(parr)[["sex"]] )
+    {
+    matplot( A.pt, crv <- parr[mod,"T1",sx,,wh,1]/
+                          (parr[mod,"T1",sx,,wh,1]+
+                           parr[mod,"T2",sx,,wh,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) )
+    text( rep(50,5), crv["50",], wh, cex=0.5, adj=1 )
+    }
+ mtext( "Age", side=1, line=2, outer=TRUE )
+ }
> prfr( "glm" )


> prfr( "gam" )
```

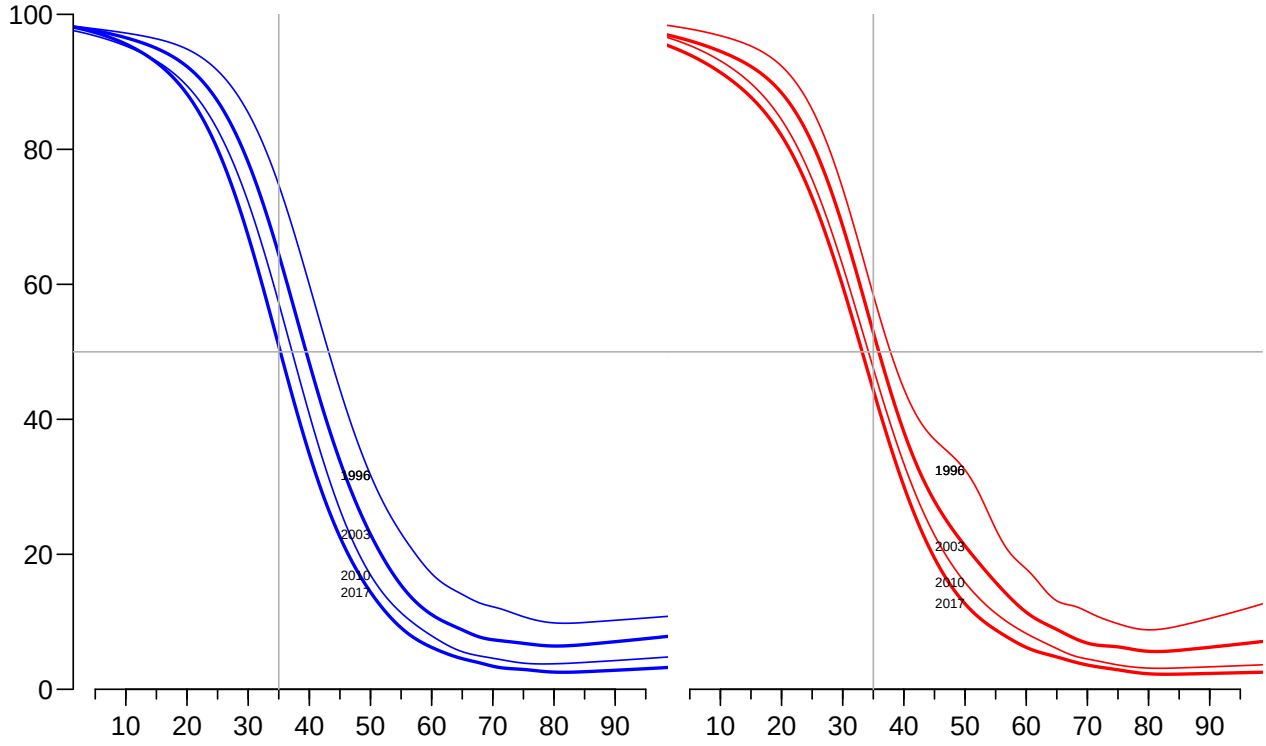


Figure 4.11: *Fraction of T1D among all diabetes patients for dates 1996-01-01,2003-01-01 etc., based on* `glm` *models. In both panels, at age 50 the upper curve is for 1997, and the lower for 2017, reflecting the steeper increase in T2D.* `./graph/prev-p5-ratio-glm`

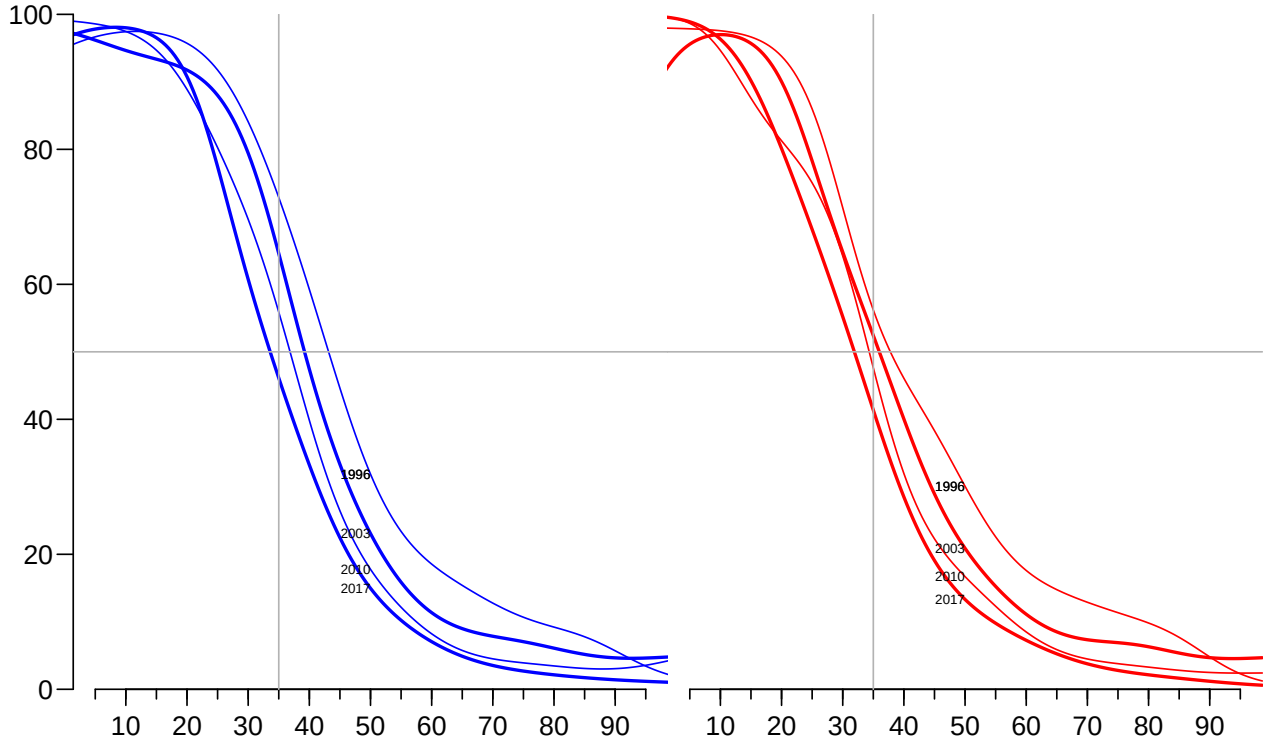


Figure 4.12: *Fraction of T1D among all diabetes patients for dates 1996-01-01,2003-01-01 etc., based on* `gam` *models. In both panels, at age 50 the upper curve is for 1997, and the lower for 2016, reflecting the steeper increase in T2D.* `./graph/prev-p5-ratio-gam`

4.4 Consistency of population size

For the prediction of future prevalent *number* of diabetes patients, we would like to see how the derived numbers from the registers match with the forecast numbers from DST:

```
> str( prN )
'data.frame':        22000 obs. of  10 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","W": 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 2 3 2 0 3 6 1 3 ...
 $ T2 : num  0 0 0 0 0 0 0 0 0 0 ...
 $ nD : num  3822 3939 3930 3872 3685 ...
 $ N  : num  3822 3939 3932 3875 3687 ...
 $ DM : num  0 0 2 3 2 0 3 6 1 3 ...
 $ X  : num  0 0 2 3 2 0 3 6 1 3 ...
> rbef <- aggregate( prN$N, prN[,c("sex","A","P")], FUN=sum )
> names( rbef )[4] <- "N"
> str( rbef )
'data.frame':        4400 obs. of  4 variables:
 $ sex: Factor w/ 2 levels "M","W": 1 2 1 2 1 2 1 2 1 2 ...
 $ A  : num  0 0 1 1 2 2 3 3 4 4 ...
 $ P  : num  1996 1996 1996 1996 1996 ...
 $ N  : num  36271 34214 36169 34567 35124 ...
> load( "../nydata/pop.Rda" )
> beff$sex <- Relevel( beff$sex, 2:1 )
> str( beff )
'data.frame':        4800 obs. of  4 variables:
 $ sex: Factor w/ 2 levels "M","W": 2 1 2 1 2 1 2 1 2 1 ...
 $ A  : int  0 0 1 1 2 2 3 3 4 4 ...
 $ P  : int  2017 2017 2017 2017 2017 2017 2017 2017 2017 2017 ...
 $ N  : int  30054 31970 28889 30364 28451 30069 28302 29674 29187 31101 ...
 - attr(*, "Label")= chr "DK population forecasts as of 2017 - 2040-01-01"
```

We then plot the cumulative number of persons to a given age as a function of calendar time, and observe if there are substantial jumps at 2017:

```
> rdat <- xtabs( N ~ sex + P + A, data=subset(rbef,A<100) )
> fdat <- xtabs( N ~ sex + P + A, data=beff )
> rdat <- apply( rdat, 2:1, cumsum )
> fdat <- apply( fdat, 2:1, cumsum )
> fCtable( fdat[c(1,11,21)+70,,], col.vars=c(1,3) )
     A          70                  80                  90
     sex         M         W         M         W         M         W
P
2017     2,555,878 2,501,585 2,778,083 2,752,516 2,851,241 2,863,055
2018     2,558,874 2,502,291 2,794,506 2,766,317 2,870,579 2,878,896
2019     2,560,821 2,502,383 2,807,749 2,777,491 2,887,283 2,893,184
2020     2,563,475 2,503,978 2,818,662 2,787,456 2,902,065 2,906,534
2021     2,567,258 2,506,856 2,828,368 2,796,885 2,915,671 2,919,545
2022     2,570,680 2,509,624 2,836,943 2,805,211 2,928,639 2,932,238
2023     2,574,862 2,514,078 2,845,028 2,813,316 2,941,290 2,944,920
2024     2,579,071 2,518,306 2,851,115 2,819,365 2,953,745 2,957,489
2025     2,583,050 2,522,302 2,856,448 2,824,171 2,965,993 2,969,950
```

```
2026     2,587,861 2,526,897 2,860,109 2,827,342 2,977,829 2,981,946
2027     2,591,869 2,531,299 2,862,602 2,829,478 2,989,246 2,993,506
2028     2,595,651 2,535,157 2,864,430 2,830,768 3,000,146 3,004,375
2029     2,599,579 2,539,280 2,867,285 2,832,811 3,010,487 3,014,453
2030     2,603,205 2,543,038 2,871,708 2,836,600 3,020,232 3,024,017
2031     2,607,020 2,546,856 2,877,533 2,841,500 3,029,560 3,033,083
2032     2,609,258 2,549,302 2,883,167 2,846,179 3,038,262 3,041,331
2033     2,611,521 2,551,481 2,889,166 2,851,807 3,046,481 3,049,024
2034     2,612,743 2,552,648 2,894,871 2,856,789 3,053,712 3,055,331
2035     2,611,894 2,551,657 2,900,013 2,861,191 3,060,400 3,060,741
2036     2,609,970 2,549,954 2,905,510 2,865,707 3,066,345 3,065,112
2037     2,607,214 2,546,928 2,910,074 2,869,739 3,071,716 3,068,794
2038     2,603,001 2,542,496 2,914,186 2,873,100 3,076,755 3,071,971
2039     2,601,496 2,540,524 2,918,245 2,876,457 3,082,133 3,075,405
2040     2,601,961 2,540,965 2,921,827 2,879,336 3,087,915 3,079,548
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> matplot( c( as.numeric(dimnames(rdat)[[2]]),
+             as.numeric(dimnames(fdat)[[2]]) ),
+          t( cbind( rdat[1:20*5,,"M"],
+                    fdat[1:20*5,,"M"] ) )/1000,
+          type="l", lty=1, lwd=2, col="blue",
+          xlab="Date", ylab="Cumulative population size by age" )
> matlines( c( as.numeric(dimnames(rdat)[[2]]),
+              as.numeric(dimnames(fdat)[[2]]) ),
+          t( cbind( rdat[1:20*5,,"W"],
+                    fdat[1:20*5,,"W"] ) )/1000,
+          type="l", lty=1, lwd=2, col="red" )
> abline( v=2016.5+0:1, col=gray(0.7) )


----------------------
2019-01-06 at 11:06:30
Time elapsed: 00:04:07
----------------------
```

. . . now input from `rates.tex`

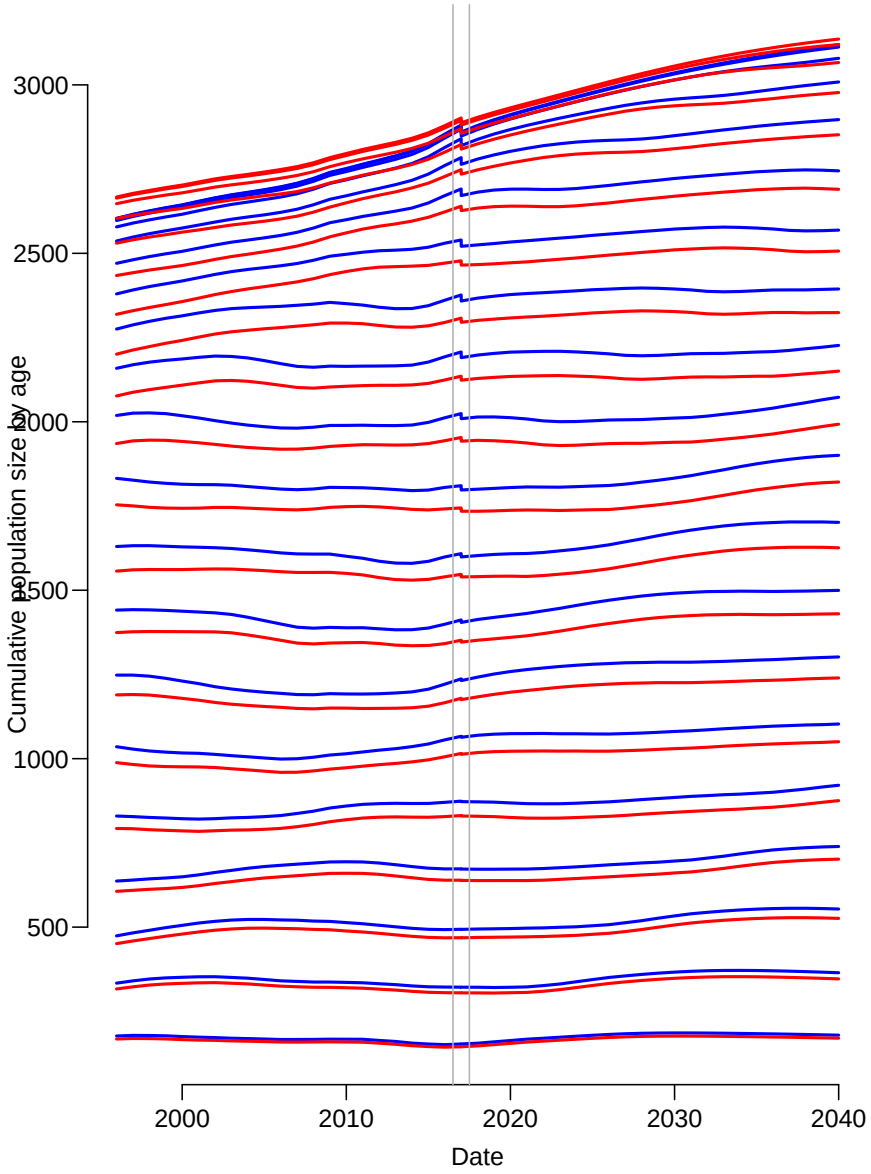


Figure 4.13: *Cumulative number of persons in ages 0,5,10... as a function of date. Note that up till quite late, the cumulative number of men is larger than that of women. This is partly due to the natural birth surplus of men, but mainly because the immigrants are predominantly men.* `./graph/prev-cumpop`

Chapter 5

Rate data

We read the register-derived datasets — the `rtL` is the dataset with only Late diagnoses of DM, that is from 1996-01-01, so the part of the follow-up where we are reasonably sure about the date of diagnosis (even in the absence of the DADD information).

```
> rt  <- read_sas("../nydata/FUtot.sas7bdat")  ; str(rt)
Classes 'tbl_df', 'tbl' and 'data.frame':        360946 obs. of  14 variables:
 $ sex  : num  1 1 1 1 1 1 1 1 1 1 ...
 $ state: chr  "T1" "T1" "T1" "T1" ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
  ..- attr(*, "label")= chr "Age at FU"
 $ P    : num  1997 1999 1999 2000 2000 ...
  ..- attr(*, "label")= chr "Date of FU"
 $ C    : num  1996 1998 1999 1999 1999 ...
  ..- attr(*, "label")= chr "Date of birth"
 $ dur  : num  0 0 0 0 0.2 0 0 0 0.2 0 ...
  ..- attr(*, "label")= chr "DMdur at FU"
 $ Dcvd : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dcan : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dres : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Doth : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y    : num  1.43e-04 2.26e-05 7.67e-05 3.79e-04 1.76e-04 ...
  ..- attr(*, "label")= chr "PY (1000s)"
 $ T1   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "label")= chr "FUTOT"
> rtL <- read_sas("../nydata/FUtotL.sas7bdat") ; str(rtL)
Classes 'tbl_df', 'tbl' and 'data.frame':        184079 obs. of  14 variables:
 $ sex  : num  1 1 1 1 1 1 1 1 1 1 ...
 $ state: chr  "T1" "T1" "T1" "T1" ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
  ..- attr(*, "label")= chr "Age at FU"
 $ P    : num  1997 1999 1999 2000 2000 ...
  ..- attr(*, "label")= chr "Date of FU"
 $ C    : num  1996 1998 1999 1999 1999 ...
  ..- attr(*, "label")= chr "Date of birth"
 $ dur  : num  0 0 0 0 0.2 0 0 0 0.2 0 ...
  ..- attr(*, "label")= chr "DMdur at FU"
 $ Dcvd : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dcan : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dres : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
 $ Doth : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y    : num  1.43e-04 2.26e-05 7.67e-05 3.79e-04 1.76e-04 ...
  ..- attr(*, "label")= chr "PY (1000s)"
 $ T1   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "label")= chr "FUTOTL"
```

The datasets are primarily classified by `state`, the state *from* which events (`T1`, `T2` or `D`) may occur. Since 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, but only for teh states T1 and T2 — the variable has no quantitative meaning for noDM, so for theres it is kept as 0:

```
> 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","W") ),
+               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       .       5       .       2       .       2
 0.67       5       .       6       .       2       .
 1.33       .       8       .      11       .       7
 1.67      11       .       9       .      12       .
 2.33       .      13       .      10       .      13
 2.67      11       .      13       .      11       .
 3.33       .      16       .      15       .      11
 3.67      15       .      14       .      17       .
 4.33       .      16       .      17       .      19
 4.67      14       .      15       .      19       .
```

Note that this recoding means that `dur` is never 0 for T1 and T2 persons, only for noDM.

```
> tt <- addmargins( xtabs( cbind( T1, T2, D, Y ) ~ state + dur, data=rt ), 2 )
> fCtable( tt[,c(dim(tt)[2],1:4),],row.vars=1:2, w=7 )
                T1      T2       D       Y
state dur
noDM  Sum   19,712 343,952 988,569 111,456
      0     19,712 343,952 988,569 111,456
      0.1        .       .       .       .
      0.35       .       .       .       .
      0.75       .       .       .       .
T1    Sum        .       .  12,762     501
      0          .       .       .       .
      0.1        .       .     159       4
      0.35       .       .     197       6
      0.75       .       .     260      10
T2    Sum        .       . 149,000   3,096
      0          .       .       .       .
```

```
      0.1         .       .   4,088      68
      0.35        .       .   4,527     101
      0.75        .       .   6,690     168
> with( rt, table(P-A-C) )
     0      1
177522 176828
```

We do 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","W") ),
+                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       .       5       .       2       .       2
  0.67       5       .       6       .       2       .
  1.33       .       8       .      11       .       7
  1.67       7       .       9       .      12       .
  2.33       .       8       .      10       .      13
  2.67       5       .      10       .      11       .
  3.33       .      10       .      10       .      11
  3.67       7       .      10       .      14       .
  4.33       .       8       .      12       .      14
  4.67       6       .      10       .      15       .


> tt <- addmargins( xtabs( cbind( T1, T2, D, Y ) ~ state + dur, data=rtL ), 2 )
> fCtable( tt[,c(dim(tt)[2],1:4),],row.vars=1:2, w=7 )
                 T1      T2       D       Y
state dur
noDM  Sum    19,712 343,952 988,569 111,456
      0      19,712 343,952 988,569 111,456
      0.1         .       .       .       .
      0.35        .       .       .       .
      0.75        .       .       .       .
T1    Sum         .       .   3,528     184
      0           .       .       .       .
      0.1         .       .     158       4
      0.35        .       .     193       6
      0.75        .       .     239       9
T2    Sum         .       . 100,906   2,456
      0           .       .       .       .
      0.1         .       .   4,062      68
      0.35        .       .   4,433      99
      0.75        .       .   6,063     159
> with( rt, table(P-A-C) )
     0      1
177522 176828
```

Finally we save the groomed dataframes for analysis:

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

```
----------------------
2019-01-06 at 11:58:54
Time elapsed: 00:00:06
----------------------
```

. . . now input from `inc.tex`

Chapter 6

DM incidence

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

```
> library( Epi )
> library( mgcv )
> start()
---------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-01-10 11:22:06
---------------------------------------

> load( file="../nydata/inits.Rda" )
> load( file="../nydata/rt.Rda" )
```

Modeling incidence rates of diabetes will involve only persons in the `noDM` state, which is the same in the two datasets.

Hence, we restrict 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':        8400 obs. of  17 variables:
 $ sex  : Factor w/ 2 levels "M","W": 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 ...
 $ Dcvd : num  0 0 0 0 1 0 1 0 0 0 ...
 $ Dcan : num  0 0 1 0 1 0 1 0 1 0 ...
 $ Dres : num  3 0 0 0 1 1 0 0 1 0 ...
 $ Doth : num  25 0 17 1 15 1 9 0 19 0 ...
 $ Y    : num  17.9 17.6 17.3 17.7 17.1 ...
 $ T1   : num  0 0 1 0 0 0 1 1 2 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 2 0 ...
 $ D    : num  28 0 18 1 18 2 11 0 21 0 ...
 $ 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=1:3 ),
+          col.vars=3:2, w=7 )
```

```
             T1                      T2                      Sum
     sex      M       W     Sum       M       W     Sum       M       W     Sum
P
1996         678     516   1,194   6,115   5,290  11,405   6,793   5,806  12,599
1997         684     489   1,173   5,839   4,918  10,757   6,523   5,407  11,930
1998         657     454   1,111   6,529   5,295  11,824   7,186   5,749  12,935
1999         592     413   1,005   6,739   5,707  12,446   7,331   6,120  13,451
2000         596     392     988   6,593   5,604  12,197   7,189   5,996  13,185
2001         586     415   1,001   6,795   5,449  12,244   7,381   5,864  13,245
2002         602     386     988   8,022   7,334  15,356   8,624   7,720  16,344
2003         545     386     931   9,146   7,673  16,819   9,691   8,059  17,750
2004         509     388     897   9,259   7,751  17,010   9,768   8,139  17,907
2005         517     379     896   8,174   6,510  14,684   8,691   6,889  15,580
2006         554     382     936   8,172   5,940  14,112   8,726   6,322  15,048
2007         564     384     948   8,738   6,792  15,530   9,302   7,176  16,478
2008         546     367     913   9,846   7,554  17,400  10,392   7,921  18,313
2009         568     357     925  10,762   7,720  18,482  11,330   8,077  19,407
2010         529     367     896  11,867   8,704  20,571  12,396   9,071  21,467
2011         496     358     854  15,593  13,150  28,743  16,089  13,508  29,597
2012         486     315     801  12,782  10,017  22,799  13,268  10,332  23,600
2013         471     351     822  10,215   7,971  18,186  10,686   8,322  19,008
2014         465     341     806   9,883   7,358  17,241  10,348   7,699  18,047
2015         476     375     851   9,987   7,638  17,625  10,463   8,013  18,476
2016         460     316     776  10,666   7,855  18,521  11,126   8,171  19,297
Sum       11,581   8,131  19,712 191,722 152,230 343,952 203,303 160,361 363,664
```

```
> fCtable( addmargins( xtabs( cbind(T1,T2,Y) ~ P + sex, data=st  ),
+                      margin=2 ),
+          col.vars=3:2, w=6 )
```

```
             T1                    T2                    Y
     sex      M      W    Sum       M      W    Sum      M      W    Sum
P
1996         678    516  1,194   6,115  5,290 11,405  2,565  2,629  5,194
1997         684    489  1,173   5,839  4,918 10,757  2,572  2,637  5,209
1998         657    454  1,111   6,529  5,295 11,824  2,578  2,643  5,221
1999         592    413  1,005   6,739  5,707 12,446  2,583  2,647  5,230
2000         596    392    988   6,593  5,604 12,197  2,589  2,653  5,241
2001         586    415  1,001   6,795  5,449 12,244  2,595  2,659  5,255
2002         602    386    988   8,022  7,334 15,356  2,601  2,664  5,265
2003         545    386    931   9,146  7,673 16,819  2,604  2,666  5,270
2004         509    388    897   9,259  7,751 17,010  2,605  2,668  5,273
2005         517    379    896   8,174  6,510 14,684  2,608  2,671  5,279
2006         554    382    936   8,172  5,940 14,112  2,613  2,676  5,289
2007         564    384    948   8,738  6,792 15,530  2,621  2,683  5,304
2008         546    367    913   9,846  7,554 17,400  2,632  2,692  5,324
2009         568    357    925  10,762  7,720 18,482  2,640  2,702  5,342
2010         529    367    896  11,867  8,704 20,571  2,645  2,709  5,354
2011         496    358    854  15,593 13,150 28,743  2,648  2,713  5,362
2012         486    315    801  12,782 10,017 22,799  2,651  2,715  5,366
2013         471    351    822  10,215  7,971 18,186  2,657  2,721  5,378
2014         465    341    806   9,883  7,358 17,241  2,671  2,732  5,403
2015         476    375    851   9,987  7,638 17,625  2,691  2,747  5,438
2016         460    316    776  10,666  7,855 18,521  2,704  2,756  5,460
```

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

A broad overview of rates:

```
> stall <-
+ stat.table( index = list(P,sex),
+          contents = list( T1=ratio(T1,Y),
+                           T2=ratio(T2,Y),
+                           DM=ratio(T1+T2,Y) ),
+           margins = TRUE,
+              data = st )
> round( ftable( stall, row.vars=2 ),2 )
      contents   T1              T2              DM
      sex         M    W Total    M    W Total    M    W Total
P
1996           0.26 0.20  0.23 2.38 2.01  2.20 2.65 2.21  2.43
1997           0.27 0.19  0.23 2.27 1.87  2.07 2.54 2.05  2.29
1998           0.25 0.17  0.21 2.53 2.00  2.26 2.79 2.18  2.48
1999           0.23 0.16  0.19 2.61 2.16  2.38 2.84 2.31  2.57
2000           0.23 0.15  0.19 2.55 2.11  2.33 2.78 2.26  2.52
2001           0.23 0.16  0.19 2.62 2.05  2.33 2.84 2.21  2.52
2002           0.23 0.14  0.19 3.08 2.75  2.92 3.32 2.90  3.10
2003           0.21 0.14  0.18 3.51 2.88  3.19 3.72 3.02  3.37
2004           0.20 0.15  0.17 3.55 2.91  3.23 3.75 3.05  3.40
2005           0.20 0.14  0.17 3.13 2.44  2.78 3.33 2.58  2.95
2006           0.21 0.14  0.18 3.13 2.22  2.67 3.34 2.36  2.84
2007           0.22 0.14  0.18 3.33 2.53  2.93 3.55 2.67  3.11
2008           0.21 0.14  0.17 3.74 2.81  3.27 3.95 2.94  3.44
2009           0.22 0.13  0.17 4.08 2.86  3.46 4.29 2.99  3.63
2010           0.20 0.14  0.17 4.49 3.21  3.84 4.69 3.35  4.01
2011           0.19 0.13  0.16 5.89 4.85  5.36 6.07 4.98  5.52
2012           0.18 0.12  0.15 4.82 3.69  4.25 5.01 3.81  4.40
2013           0.18 0.13  0.15 3.84 2.93  3.38 4.02 3.06  3.53
2014           0.17 0.12  0.15 3.70 2.69  3.19 3.87 2.82  3.34
2015           0.18 0.14  0.16 3.71 2.78  3.24 3.89 2.92  3.40
2016           0.17 0.11  0.14 3.94 2.85  3.39 4.11 2.96  3.53
Total          0.21 0.14  0.18 3.48 2.70  3.09 3.69 2.84  3.26
```

We also produce an overview of the ages at inclusion:

```
> par( mfrow=c(2,3), mar=c(2,2,1,0), oma=c(2,2,0,0), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> with( rt, hist( rep(Ax,T1), breaks=0:100, col="black", ylim=c(0,600),
+                 main="M+W", ylab="", xlab="" ) )
> mtext( "T1 diabetes incident cases", side=2, line=3, las=0, cex=0.67 )
> with( subset(rt,sex=="M"),
+       hist( rep(Ax,T1), breaks=0:100, col="blue", border="blue", ylim=c(0,300),
+             main="M", xlab="", ylab="" ) )
> with( subset(rt,sex=="W"),
+       hist( rep(Ax,T1), breaks=0:100, col="red", border="red", ylim=c(0,300),
+             main="W", xlab="", ylab="" ) )
> with( rt, hist( rep(Ax,T2), breaks=0:100, col="black", ylim=c(0,12000),
+                 main="", ylab="", xlab="" ) )
> mtext( "T2 diabetes incident cases", side=2, line=3, las=0, cex=0.67 )
> with( subset(rt,sex=="M"),
+       hist( rep(Ax,T2), breaks=0:100, col="blue", border="blue", ylim=c(0,6000),
+             main="", xlab="", ylab="" ) )
> with( subset(rt,sex=="W"),
+       hist( rep(Ax,T2), breaks=0:100, col="red", border="red", ylim=c(0,6000),
+             main="", xlab="", ylab="" ) )
> mtext( rep("Age",3), at=c(1,3,5)/6, outer=TRUE, side=1, line=0, cex=0.67 )
```

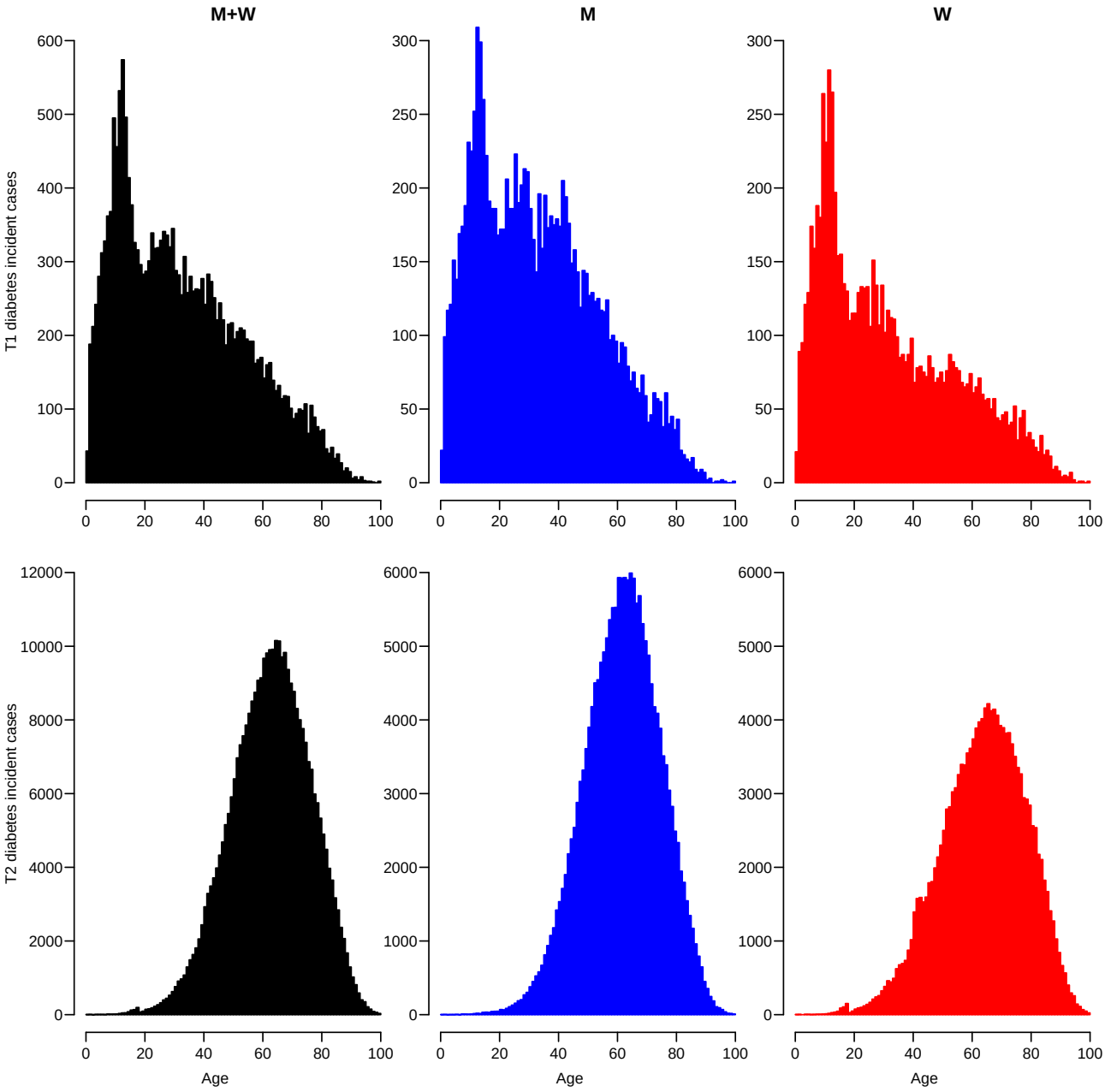


Figure 6.1: *Distribution of age at inclusion for the period 1996–2016, by sex and type of diabetes. As it stands now, there is clearly a problem around age 40 for T2 women, an effect of the PCOS/metformin definition.* `./graph/inc-inchist`

6.1 Age-Period-Cohort models

We fit APC-models for T1, T2 separately for men and women and plot the estimates together for men and women.

First we devise knots for the natural splines that we use:

```
> ( A1.kn <- with( st,    quantile( rep(    Ax, T1 ), qn(6) ) ) )
8.333333%       25% 41.66667% 58.33333%       75% 91.66667%
 7.333333 14.333333 24.333333 34.666667 47.666667 66.666667
> ( P1.kn <- with( st, c( quantile( rep( Px   , T1 ), qn(5) ), 2016 ) ) )
```

```
     10%      30%      50%      70%      90%
1997.667 2001.333 2005.667 2009.667 2014.667 2016.000
> ( C1.kn <- with( st,    quantile( rep( Px-Ax, T1 ), qn(5) ) ) )
     10%      30%      50%      70%      90%
1939.333 1961.333 1975.667 1989.667 2001.333
> ( A2.kn <- with( st,    quantile( rep(    Ax, T2 ), qn(10) ) ) )
      5%      15%      25%      35%      45%      55%      65%      75%      85%      95%
39.33333 48.33333 53.33333 57.66667 61.33333 64.66667 68.33333 72.33333 76.66667 83.66667
> ( P2.kn <- with( st, c( quantile( rep( Px   , T2 ), qn(6) ), 2016 ) ) )
8.333333%       25% 41.66667% 58.33333%       75% 91.66667%
 1998.667  2002.667  2006.667  2010.333  2012.333  2015.333  2016.000
> ( C2.kn <- with( st,    quantile( rep( Px-Ax, T2 ), qn(10) ) ) )
      5%      15%      25%      35%      45%      55%      65%      75%      85%      95%
1920.667 1929.333 1934.667 1939.333 1943.333 1946.333 1950.333 1954.667 1960.667 1970.333
```

With these knots (similar for men and women) we can fit 4 different APC-models — for men and women and T1 and T2 separately:

```
> M1 <- apc.fit( transform( subset(st,sex=="M"), A=Ax, P=Px, D=T1 )[,c("A","P","D","Y")],
+                parm="APC", ref.p=2015, npar=list(A=A1.kn,P=P1.kn,C=C1.kn), dr.extr="y" )
[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                    4194     5333.0
Age-drift              4193     5158.9  1   174.09 < 2.2e-16
Age-Cohort             4190     4516.6  3   642.26 < 2.2e-16
Age-Period-Cohort      4186     4486.8  4    29.82 5.319e-06
Age-Period             4189     5143.4 -3  -656.55 < 2.2e-16
Age-drift              4193     5158.9 -4   -15.53  0.003726
> F1 <- apc.fit( transform( subset(st,sex=="W"), A=Ax, P=Px, D=T1 )[,c("A","P","D","Y")],
+                parm="APC", ref.p=2015, npar=list(A=A1.kn,P=P1.kn,C=C1.kn), dr.extr="y"  )
[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                    4194     5070.7
Age-drift              4193     4945.2  1   125.54 < 2.2e-16
Age-Cohort             4190     4503.7  3   441.43 < 2.2e-16
Age-Period-Cohort      4186     4496.2  4     7.48  0.112484
Age-Period             4189     4927.7 -3  -431.48 < 2.2e-16
Age-drift              4193     4945.2 -4   -17.43  0.001593
> M2 <- apc.fit( transform( subset(st,sex=="M"), A=Ax, P=Px, D=T2 )[,c("A","P","D","Y")],
+                parm="APC", ref.p=2015, npar=list(A=A2.kn,P=P2.kn,C=C2.kn), dr.extr="y" )
[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                    4190    14110.8
Age-drift              4189    10087.9  1   4023.0 < 2.2e-16
```

```
Age-Cohort             4181     9911.8  8    176.1 < 2.2e-16
Age-Period-Cohort      4176     6301.7  5   3610.1 < 2.2e-16
Age-Period             4184     6627.0 -8   -325.3 < 2.2e-16
Age-drift              4189    10087.9 -5  -3460.8 < 2.2e-16
```

```
> F2 <- apc.fit( transform( subset(st,sex=="W"), A=Ax, P=Px, D=T2 )[,c("A","P","D","Y")],
+                parm="APC", ref.p=2015, npar=list(A=A2.kn,P=P2.kn,C=C2.kn), dr.extr="y" )
[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                    4190    13837.7
Age-drift              4189    11446.5  1   2391.2 < 2.2e-16
Age-Cohort             4181    10933.7  8    512.8 < 2.2e-16
Age-Period-Cohort      4176     7335.1  5   3598.6 < 2.2e-16
Age-Period             4184     8101.2 -8   -766.1 < 2.2e-16
Age-drift              4189    11446.5 -5  -3345.3 < 2.2e-16
```

When we stick to the period-major parametrization as in figure 6.2, the age-specific rates are essentially cross-sectional:

```
> apcall <- function() {
+ 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, bty="o" )
+ 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)/10,1,2,5,10,20)/10,
+            a.tic=seq(0,90,10), cp.tic=seq(1900,2020,10),
+            r.tic=c(1:10/100,1:10/10,1:10,c(1.5,2,3)*10)/10, rr.ref=0.02,
+            gap=17, r.txt="T1 diabetes incidence rate per 1000 PY", side=c(2,4) )
+ apc.lines( M1, col="black", lwd=1, knots=TRUE )
+     matshade( M1$Age[,1], M1$Age[,-1], lwd=2, col="blue" )
+  pc.matshade( M1$Per[,1], M1$Per[,-1], lwd=2, col="blue" )
+  pc.matshade( M1$Coh[,1], M1$Coh[,-1], lwd=2, col="blue" )
+     matshade( F1$Age[,1], F1$Age[,-1], lwd=2, col="red" )
+  pc.matshade( F1$Per[,1], F1$Per[,-1], lwd=2, col="red" )
+  pc.matshade( F1$Coh[,1], F1$Coh[,-1], lwd=2, 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)/10,1,2,5,10,20),
+            a.tic=seq(0,90,10), cp.tic=seq(1900,2020,10),
+            r.tic=c(1:10/100,1:10/10,1:10,c(1.5,2,3)*10), rr.ref=0.2,
+            gap=17, r.txt="T2 diabetes incidence rate per 1000 PY", side=c(1,2,4) )
+ apc.lines( M2, col="black", lwd=1, knots=TRUE )
+     matshade( M2$Age[,1], M2$Age[,-1], lwd=2, col="blue" )
+  pc.matshade( M2$Per[,1], M2$Per[,-1], lwd=2, col="blue" )
+  pc.matshade( M2$Coh[,1], M2$Coh[,-1], lwd=2, col="blue" )
+     matshade( F2$Age[,1], F2$Age[,-1], lwd=2, col="red" )
+  pc.matshade( F2$Per[,1], F2$Per[,-1], lwd=2, col="red" )
+  pc.matshade( F2$Coh[,1], F2$Coh[,-1], lwd=2, col="red" ) }
> apcall()
```

We devise a special version for the ESM, with slightly narrower age and cohort ranges plotted, and with annotation:

```
> 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, lend="butt", bty="o" )
> wa <- M1$Age[,1]<70
> wc <- M1$Coh[,1]<2005 & M1$Coh[,1]>1915
> apc.frame( a.lab=seq(0,90,20), cp.lab=seq(1920,2020,30),
```

```
+            r.lab=c(c(1,2,5)/10,1,2,5,10)/10,
+            a.tic=seq(0,90,10), cp.tic=seq(1920,2020,10),
+            r.tic=c(1:10/10,1:10,15)/10, rr.ref=0.1,
+            gap=7, r.txt="T1 diabetes incidence rate per 1000 PY", side=c(2,4) )
>    matshade( M1$Age[wa,1], M1$Age[wa,-1], lwd=2, col="blue" )
> pc.matshade( M1$Per[  ,1], M1$Per[  ,-1], lwd=2, col="blue" )
> pc.matshade( M1$Coh[wc,1], M1$Coh[wc,-1], lwd=2, col="blue", lty="22")
>    matshade( F1$Age[wa,1], F1$Age[wa,-1], lwd=2, col="red" )
> pc.matshade( F1$Per[  ,1], F1$Per[  ,-1], lwd=2, col="red" )
> pc.matshade( F1$Coh[wc,1], F1$Coh[wc,-1], lwd=2, col="red", lty="22")
> pc.points( 2015, 1, pch=16, cex=0.6, col="white" )
> pc.points( 2015, 1, pch=1 , cex=0.6 )
> text( 5, sqrt(1.5), "a", font=2, cex=1.2 )
> wa <- M1$Age[,1]>20 & M1$Age[,1]<90
> apc.frame( a.lab=seq(0,90,20), cp.lab=seq(1920,2020,30),
+            r.lab=c(c(1,2,5)/10,1,2,5,10),
+            a.tic=seq(0,90,10), cp.tic=seq(1920,2020,10),
+            r.tic=c(1:10/10,1:10,15), rr.ref=1,
+            gap=7, r.txt="T2 diabetes incidence rate per 1000 PY", side=c(1,2,4) )
>    matshade( M2$Age[wa,1], M2$Age[wa,-1], lwd=2, col="blue" )
> pc.matshade( M2$Per[  ,1], M2$Per[  ,-1], lwd=2, col="blue" )
> pc.matshade( M2$Coh[wc,1], M2$Coh[wc,-1], lwd=2, col="blue", lty="22")
>    matshade( F2$Age[wa,1], F2$Age[wa,-1], lwd=2, col="red" )
> pc.matshade( F2$Per[  ,1], F2$Per[  ,-1], lwd=2, col="red" )
> pc.matshade( F2$Coh[wc,1], F2$Coh[wc,-1], lwd=2, col="red", lty="22")
> pc.points( 2015, 1, pch=16, cex=0.6, col="white" )
> pc.points( 2015, 1, pch=1 , cex=0.6 )
> text( 5, 10*sqrt(1.5), "b", font=2, cex=1.2 )
```

We also devise the ACP-parametrization of the APC model:

```
> M1 <- apc.fit( transform( subset(st,sex=="M"), A=Ax, P=Px, D=T1 )[,c("A","P","D","Y")],
+               parm="ACP", ref.c=1980, npar=list(A=A1.kn,P=P1.kn,C=C1.kn),
+               print.AOV=FALSE, dr.extr="y" )
> F1 <- apc.fit( transform( subset(st,sex=="W"), A=Ax, P=Px, D=T1 )[,c("A","P","D","Y")],
+               parm="ACP", ref.c=1980, npar=list(A=A1.kn,P=P1.kn,C=C1.kn),
+               print.AOV=FALSE, dr.extr="y" )
> M2 <- apc.fit( transform( subset(st,sex=="M"), A=Ax, P=Px, D=T2 )[,c("A","P","D","Y")],
+               parm="ACP", ref.c=1980, npar=list(A=A2.kn,P=P2.kn,C=C2.kn),
+               print.AOV=FALSE, dr.extr="y" )
> F2 <- apc.fit( transform( subset(st,sex=="W"), A=Ax, P=Px, D=T2 )[,c("A","P","D","Y")],
+               parm="ACP", ref.c=1980, npar=list(A=A2.kn,P=P2.kn,C=C2.kn),
+               print.AOV=FALSE, dr.extr="y" )


> 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 )
> apcall()
```

6.1.1 Time-trends in rates

We can get a raw overview of the average trend in incidence, although the period/cohort effects for type 2 in figures 6.2 and 6.2 are very curved, so a linear average is not worth much for type 2. We report the average drift as % per year:

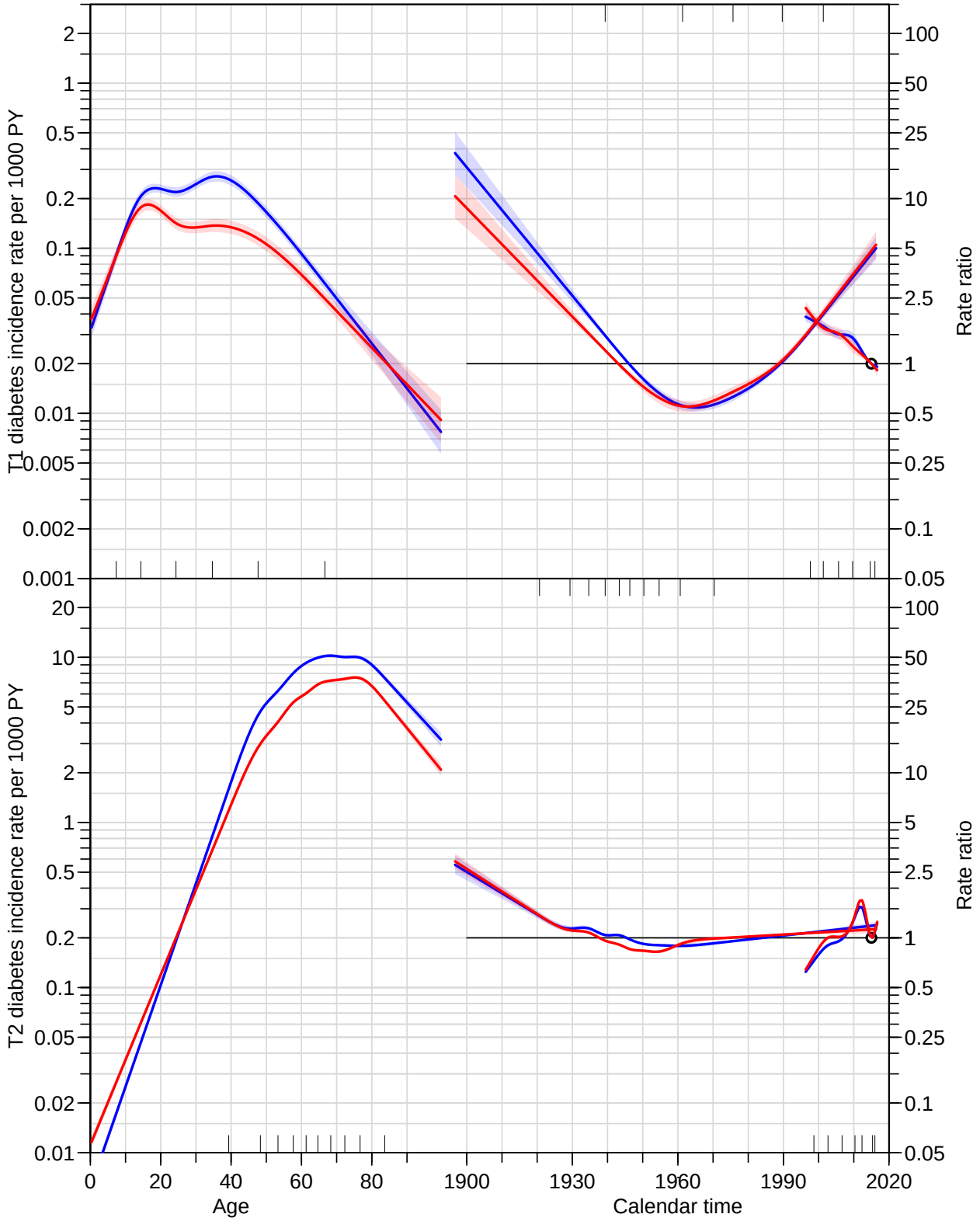


Figure 6.2: *Estimated APC-effects for T1 and T2 diabetes for men (blue) and women (red), using the APC-parametrization — age-specific rates interpretable as cross-sectional for the year 2010.*
`./graph/inc-apc`

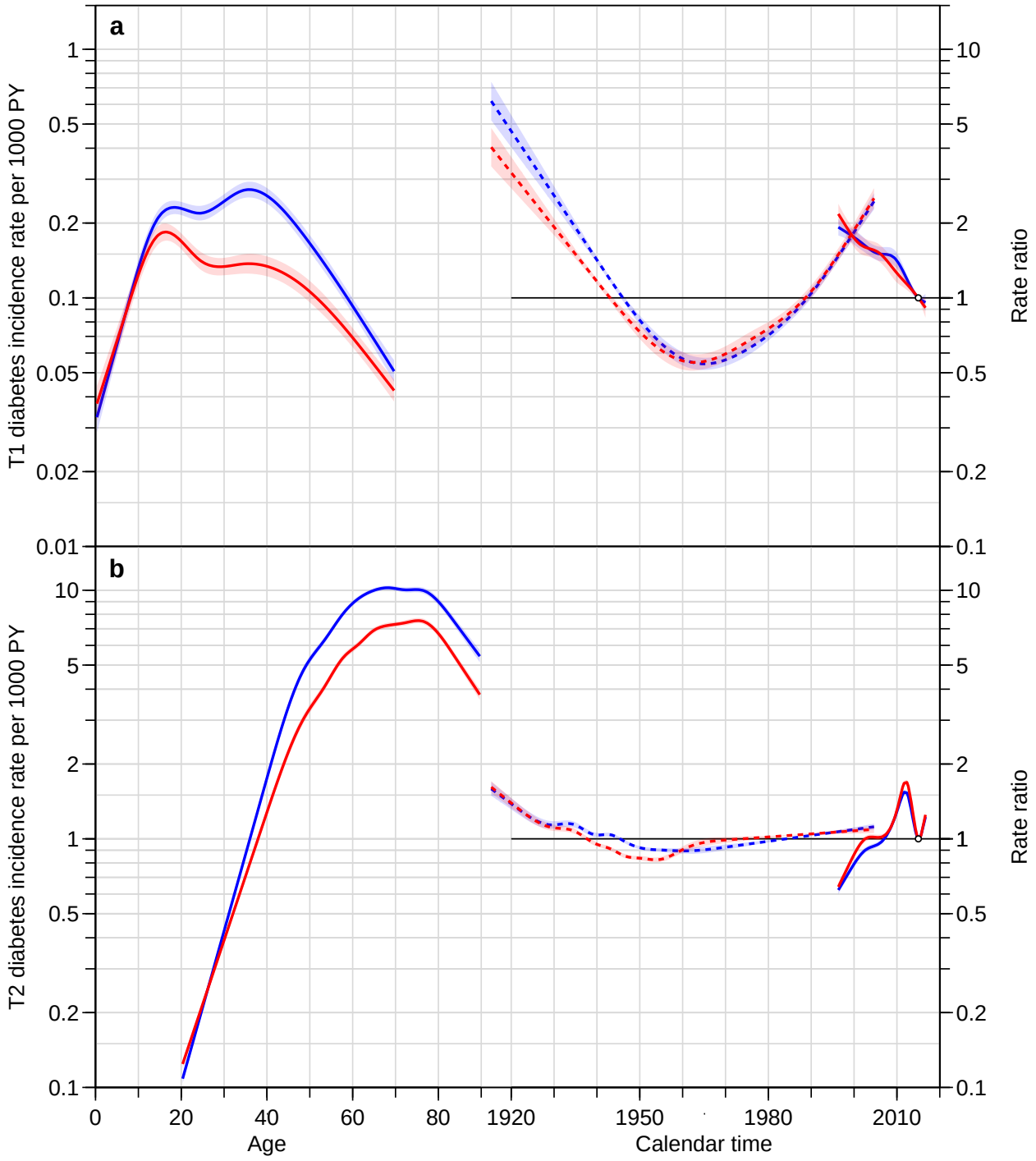


Figure 6.3: *Estimated APC-effects for T1 and T2 diabetes for men (blue) and women (red), using the APC-parametrization — age-specific rates interpretable as cross-sectional for the year 2010. Note that the scale for incidences of T1D and T2D are different, but have the same* relative *extent.* `./graph/inc-art-apc`

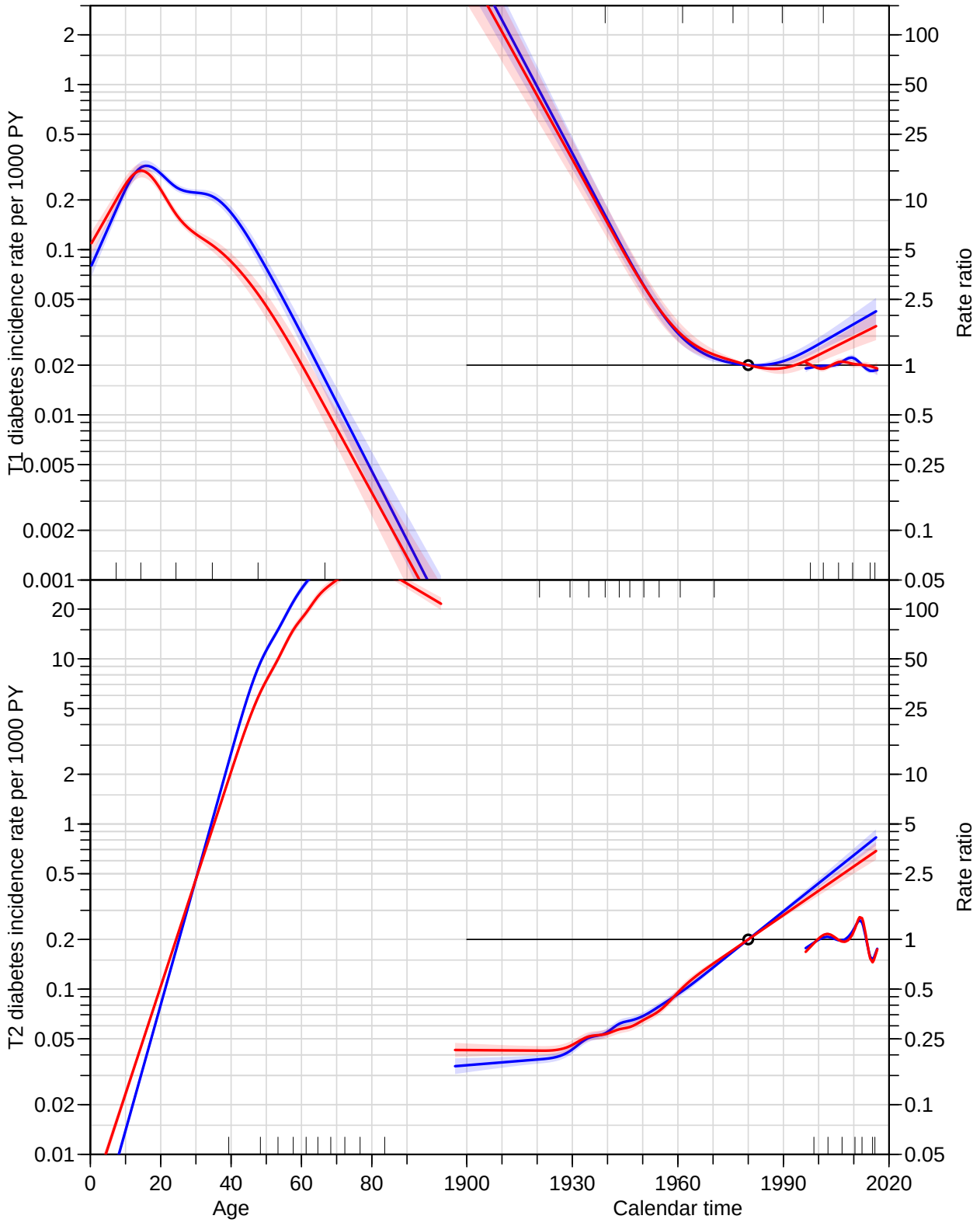


Figure 6.4: *Estimated APC-effects for T1 and T2 diabetes for men (blue) and women (red), using the ACP-parametrization — age-specific rates interpretable as longitudinal rates. The fitted values from these models are the same as for the APC-parametrization, Age-specific rates are interpretable as longitudinal rates for persons born at 1980.* `./graph/inc-acp`

```
> dr <- rbind( cbind( M1$Drift, M2$Drift )[1,],
+              cbind( F1$Drift, F2$Drift )[1,] )
> rownames( dr ) <- c("M","W")
> colnames( dr )[c(1,4)] <- c("T1","     T2")
> round( (dr-1)*100, 2 )
     T1  2.5% 97.5%      T2 2.5% 97.5%
M -3.27 -3.59 -2.94    3.43 3.25  3.61
W -3.78 -4.19 -3.37    3.16 2.98  3.34
```

The trends extracted here have been extracted using the number of events (which equals the information about the log-rate parameters) as weights in the inner product used.

6.2 Model for incidence rates

We cannot use the returned models from the `apc` objects, since these is defined in terms-specific matrices and *not* in terms of A and P, and hence not suited for prediction.

Therefore we refit the model(s) in order to be able to produce predicted rates; we also include fits of age-period and age-cohort models, as well as a simplified variant of the Lee-Carter model, here termed the local drift model (see section below):

```
> M1apc <- glm( T1 ~ Ns(    Ax, kn=M1$Knots$Age ) +
+                    Ns( Px-Ax, kn=M1$Knots$Coh ) +
+                    Ns( Px   , kn=M1$Knots$Per ),
+               offset = log( Y ),
+               family = poisson,
+                 data = subset( rt, sex=="M" & state=="noDM" ) )
> M1ap  <- update( M1apc, . ~ . - Ns( Px-Ax, kn=M1$Knots$Coh ) )
> M1ac  <- update( M1apc, . ~ . - Ns( Px   , kn=M1$Knots$Per ) )
> # local drift model
> M1al  <- update( M1ap , . ~ . - Ns( Px   , kn=M1$Knots$Per )
+                               + Ns(    Ax, kn=M1$Knots$Age[-c(3,5)], int=TRUE ):Px )
> F1apc <- glm( T1 ~ Ns(    Ax, kn=F1$Knots$Age ) +
+                    Ns( Px-Ax, kn=F1$Knots$Coh ) +
+                    Ns( Px   , kn=F1$Knots$Per ),
+               offset = log( Y ),
+               family = poisson,
+                 data = subset( rt, sex=="W" & state=="noDM" ) )
> F1ap  <- update( F1apc, . ~ . - Ns( Px-Ax, kn=F1$Knots$Coh ) )
> F1ac  <- update( F1apc, . ~ . - Ns( Px   , kn=F1$Knots$Per ) )
> # local drift model
> F1al  <- update( F1ap , . ~ . - Ns( Px   , kn=F1$Knots$Per )
+                               + Ns(    Ax, kn=F1$Knots$Age[-c(3,5)], int=TRUE ):Px )
> M2apc <- glm( T2 ~ Ns(    Ax, kn=M2$Knots$Age ) +
+                    Ns( Px-Ax, kn=M2$Knots$Coh ) +
+                    Ns( Px   , kn=M2$Knots$Per ),
+               offset = log( Y ),
+               family = poisson,
+                 data = subset( rt, sex=="M" & state=="noDM" ) )
> M2ap  <- update( M2apc, . ~ . - Ns( Px-Ax, kn=M2$Knots$Coh ) )
> M2ac  <- update( M2apc, . ~ . - Ns( Px   , kn=M2$Knots$Per ) )
> # local drift model
> M2al  <- update( M2ap , . ~ . - Ns( Px   , kn=M2$Knots$Per )
+                               + Ns(    Ax, kn=M2$Knots$Age[-c(3,5,7,9)], int=TRUE ):Px )
> F2apc <- glm( T2 ~ Ns(    Ax, kn=F2$Knots$Age ) +
+                    Ns( Px-Ax, kn=F2$Knots$Coh ) +
```

```
+                    Ns( Px   , kn=F2$Knots$Per ),
+                offset = log( Y ),
+                family = poisson,
+                  data = subset( rt, sex=="W" & state=="noDM" ) )
> F2ap  <- update( F2apc, . ~ . - Ns( Px-Ax, kn=F2$Knots$Coh ) )
> F2ac  <- update( F2apc, . ~ . - Ns( Px   , kn=F2$Knots$Per ) )
> # local drift model
> F2al  <- update( F2ap , . ~ . - Ns( Px   , kn=F2$Knots$Per )
+                               + Ns(    Ax, kn=F2$Knots$Age[-c(3,5,7,9)], int=TRUE ):Px )
> summary( fitted( M1$Model ) - fitted( M1apc ) )
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-2.398e-14 -6.418e-17  9.770e-15  1.557e-14  2.487e-14  8.082e-14
> summary( fitted( F1$Model ) - fitted( F1apc ) )
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-4.174e-14 -1.465e-14 -5.551e-15 -8.441e-15 -6.661e-16  1.688e-14
> summary( fitted( M2$Model ) - fitted( M2apc ) )
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-1.364e-12  1.221e-15  1.066e-14  4.885e-14  8.171e-14  1.222e-12
> summary( fitted( F2$Model ) - fitted( F2apc ) )
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-1.563e-12 -7.105e-14  3.553e-15 -2.751e-14  3.553e-14  6.821e-13
```

From the summaries above we see that the models actually *are* the same as those fitted by `apc.fit`. Moreover, we can use the latter fitted models to make predictions, regardless of the over-parametrization (we will get a warning, though). Recall that the `Y` was scaled to be person-millennia, 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).

6.2.1 `gam` models

For the sake of completeness we also fit the same models using the `gam` machinery, that automatically adjusts the number and location of knots:

```
> library( mgcv )
> M1gam <- gam( T1 ~ s(Ax) + s(Px) + s(Cx),
+                offset = log( Y ),
+                family = poisson,
+                  data = subset( rt, sex=="M" & state=="noDM" ) )
> M2gam <- update( M1gam, T2 ~ . )
> F1gam <- update( M1gam, data = subset( rt, sex=="W" & state=="noDM" ) )
> F2gam <- update( F1gam, T2 ~ . )
> # for the sake of completeness we try the simplified LCa model too:
> M1lg <- update( M1gam, . ~ s(Ax) + s(Ax,by=Px) )
> F1lg <- update( F1gam, . ~ s(Ax) + s(Ax,by=Px) )
> M2lg <- update( M2gam, . ~ s(Ax) + s(Ax,by=Px) )
> F2lg <- update( F2gam, . ~ s(Ax) + s(Ax,by=Px) )
```

6.2.2 Local drifts

Above we also fitted the local-drift models; namely the one that has an age-interaction with a linear term in period. This is a model where we can address what has been called "local drift" — the period-slope, $k(a)$, at different ages:

$$\log\big(\lambda(a,p)\big) = f(a) + k(a) \times p$$

— also called a "time-varying coefficients" model.

The function $k(a)$ is constant if data follow an age-drift model, so this model is almost the simplest possible interaction model expanding the age-drift model — except of course for the corresponding cohort-model, $\log\big(\lambda(a,p)\big) = f(a) + k(a) \times (p-a)$. Here we extract the local drifts from the models (the use of 2001 and 2000 is completely immaterial, any two dates one year apart will do the job):

```
> apt <- 5:80
> nd1 <- data.frame( Ax=apt, Px=2001, Y=1 )
> nd0 <- data.frame( Ax=apt, Px=2000, Y=1 )
> ldM1 <- ci.exp( M1al, list( nd1, nd0 ) )
> ldF1 <- ci.exp( F1al, list( nd1, nd0 ) )
> ldM2 <- ci.exp( M2al, list( nd1, nd0 ) )
> ldF2 <- ci.exp( F2al, list( nd1, nd0 ) )
> lxM1 <- ci.exp( M1lg, list( nd1, nd0 ) )
> lxF1 <- ci.exp( F1lg, list( nd1, nd0 ) )
> lxM2 <- ci.exp( M2lg, list( nd1, nd0 ) )
> lxF2 <- ci.exp( F2lg, list( nd1, nd0 ) )
```

With the parameters extracted we can make the relevant plots

```
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> matshade( apt, (cbind(ldM1,ldF1,ldM2,ldF2)-1)*100, plot=TRUE,
+           col=c("blue","red"), lty=rep(c("solid","11"),each=2), lend="butt", lwd=3,
+           xlab="Age", ylab="Annual change in incidence rates (%)", ylim=c(-10,10) )
> axis( side=1, at=1:8*10, labels=NA, tcl=-0.4 )
> axis( side=1, at=2:16*5, labels=NA, tcl=-0.2 )
> abline(h=0)


> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> matshade( apt, (cbind(lxM1,lxF1,lxM2,lxF2)-1)*100, plot=TRUE, ylim=c(-10,5),
+           col=c("blue","red"), lty=rep(c("solid","11"),each=2), lend="butt", lwd=3,
+           xlab="Age", ylab="Annual change in incidence rates (%)" )
> axis( side=1, at=1:8*10, labels=NA, tcl=-0.4 )
> axis( side=1, at=2:16*5, labels=NA, tcl=-0.2 )
> abline(h=0)
```

From figure 6.5 we see that there is a consistent increase in incidence rates of T2D across all ages, largely 2–3%/year in ages 50–60, but only about 1% per year in older ages. On the other hand there is a increase in T1D incidence of about 2%/year for persons under 20 years of age, whereas there is a decrease of some 4%/year in T1D incidence rates for persons around 50.

6.3 Collecting estimated incidence rates

In order to hold predicted rates from both types of models, we set up a prediction frame with ages for 15 different cohorts, and a similar frames for predicting cross-sectional rates at different dates and rates by data for select ages:

```
> a.pt <- 1:99
> p.pt <- seq(1996,2017,0.2)
> prfrm <- data.frame( expand.grid( Ax=a.pt, Px=p.pt ), Y=1 )
> prfrm <- transform( prfrm, Cx=Px-Ax )
> prArr <- NArray( list( Ax = a.pt,
```

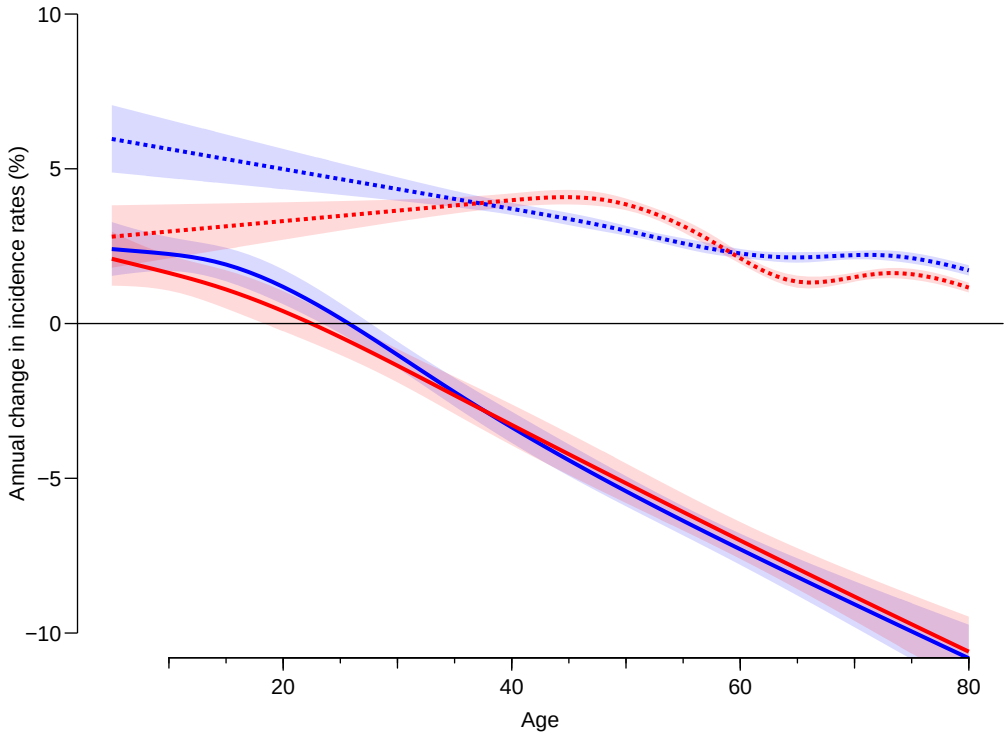


Figure 6.5: *Local drifts for T1D (full lines) and T2D (dotted lines), men in blue, women in red. Natural spline models.* `./graph/inc-ldrift`

```
+                          Px = p.pt,
+                               c("Est","lo","hi"),
+                         typ = c("T1","T2"),
+                         sex = c("M","W"),
+                         mod = c("AP","APC","gam") ) )
> prArr[,,"T1","M","AP" ] <- ci.pred( M1ap , prfrm )
> prArr[,,"T1","M","APC"] <- ci.pred( M1apc, prfrm )
> prArr[,,"T1","M","gam"] <- ci.pred( M1gam, prfrm )
> prArr[,,"T1","W","AP" ] <- ci.pred( F1ap , prfrm )
> prArr[,,"T1","W","APC"] <- ci.pred( F1apc, prfrm )
> prArr[,,"T1","W","gam"] <- ci.pred( F1gam, prfrm )
> prArr[,,"T2","M","AP" ] <- ci.pred( M2ap , prfrm )
> prArr[,,"T2","M","APC"] <- ci.pred( M2apc, prfrm )
> prArr[,,"T2","M","gam"] <- ci.pred( M2gam, prfrm )
> prArr[,,"T2","W","AP" ] <- ci.pred( F2ap , prfrm )
> prArr[,,"T2","W","APC"] <- ci.pred( F2apc, prfrm )
> prArr[,,"T2","W","gam"] <- ci.pred( F2gam, prfrm )
> length( prArr )
[1] 377784
```

6.3.1 Rates by age for different periods (cross-sectional)

Here we plot the incidence rates separately for T1 and T2, men and women showing incidence rates by age at dates 1996-01-01,. . . ,2017-01-01, 7 years apart.

```
> ( wh <- paste( seq(1996,2017,7) ) )
```

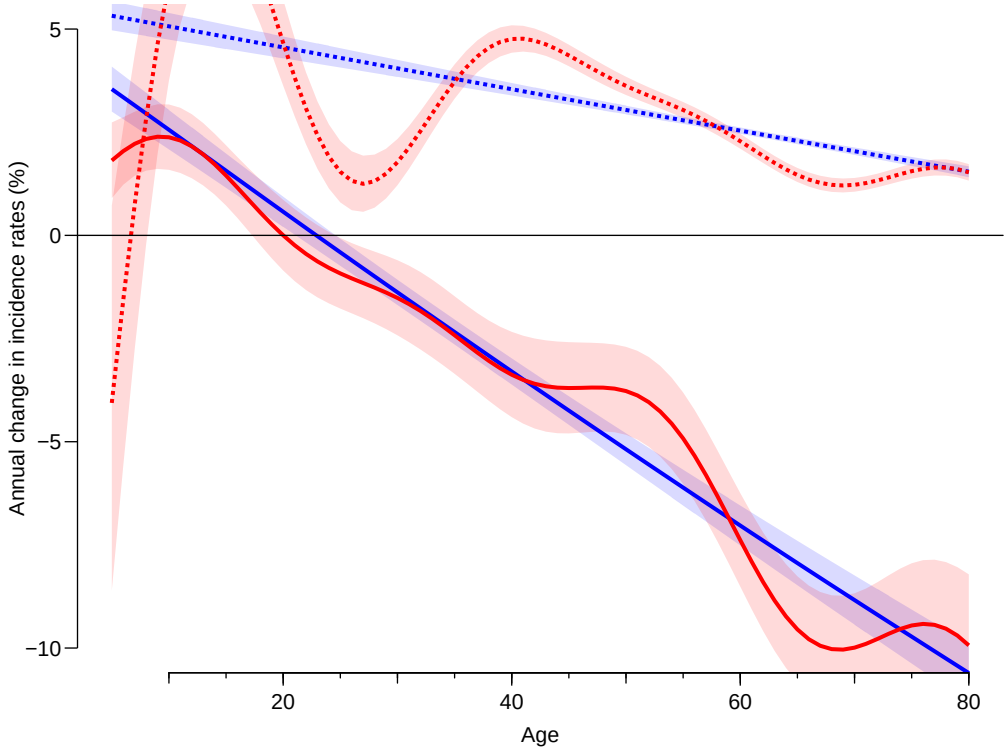


Figure 6.6: *Local drifts for T1D (full lines) and T2D (dotted lines), men in blue, women in red.* `gam` *models.* `./graph/inc-lxdrift`

```
[1] "1996" "2003" "2010" "2017"
> pl.age <- function( mod="APC" )
+    {
+ par( mfrow=c(2,2), mar=c(3,0,0,0), oma=c(0,4,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ for( tp in dimnames(prArr)[["typ"]] )
+ for( sx in dimnames(prArr)[["sex"]] )
+    {
+    matplot( a.pt, prArr[,wh,1,tp,sx,mod],
+             type="l", lwd=2, lty=1,
+             col=if(sx=="M") "blue" else "red",
+             xlab="", ylab="", xlim=c(0,95),
+             ylim=if(tp=="T1") c(0.02,1) else c(0.01,20),
+             yaxs="i", xaxt="n", yaxt="n", log="y" )
+    for( i in wh )
+       polygon( c(a.pt,rev(a.pt)), c(prArr[,i,2,tp,sx,mod],
+                                     rev(prArr[,i,3,tp,sx,mod])),
+                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( 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 incidence rate per 1000 PY"),
+                         side=2, line=2.5, 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 )
+    }
> pl.age()
```

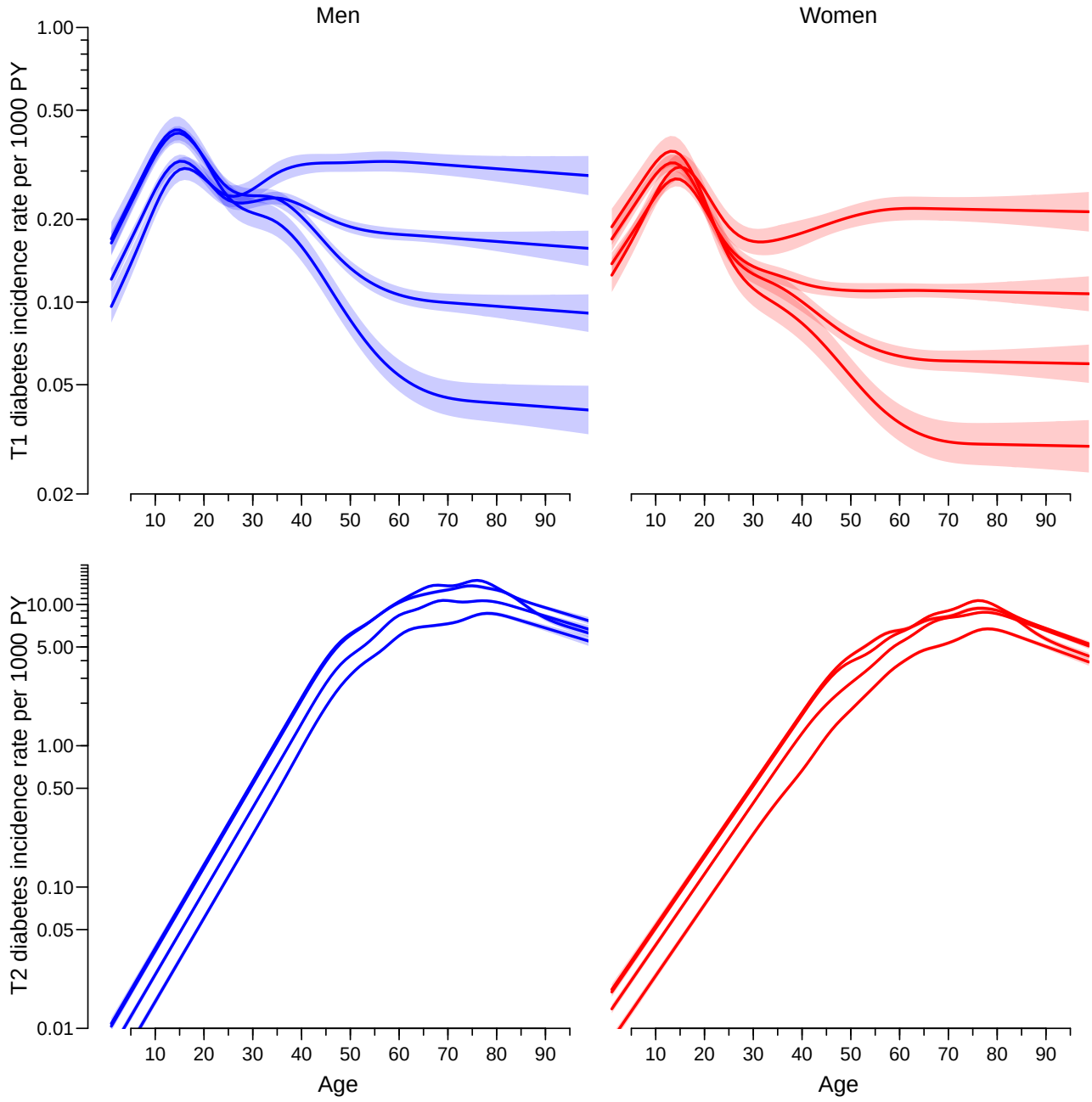


Figure 6.7: *Incidence rates of T1 and T2 diabetes in Denmark at different dates as predicted from an age-period-cohort model using* `glm`. *Note the y-axes are different between T1 and T2, with different relative extent as well. Blue: men, red: women. Shaded areas are 95% confidence intervals (almost invisible for T2 incidence).* `./graph/inc-age-apc`

Here is the same for the gam models:

```
> pl.age("gam")
```

```
> pl.age("AP")
```

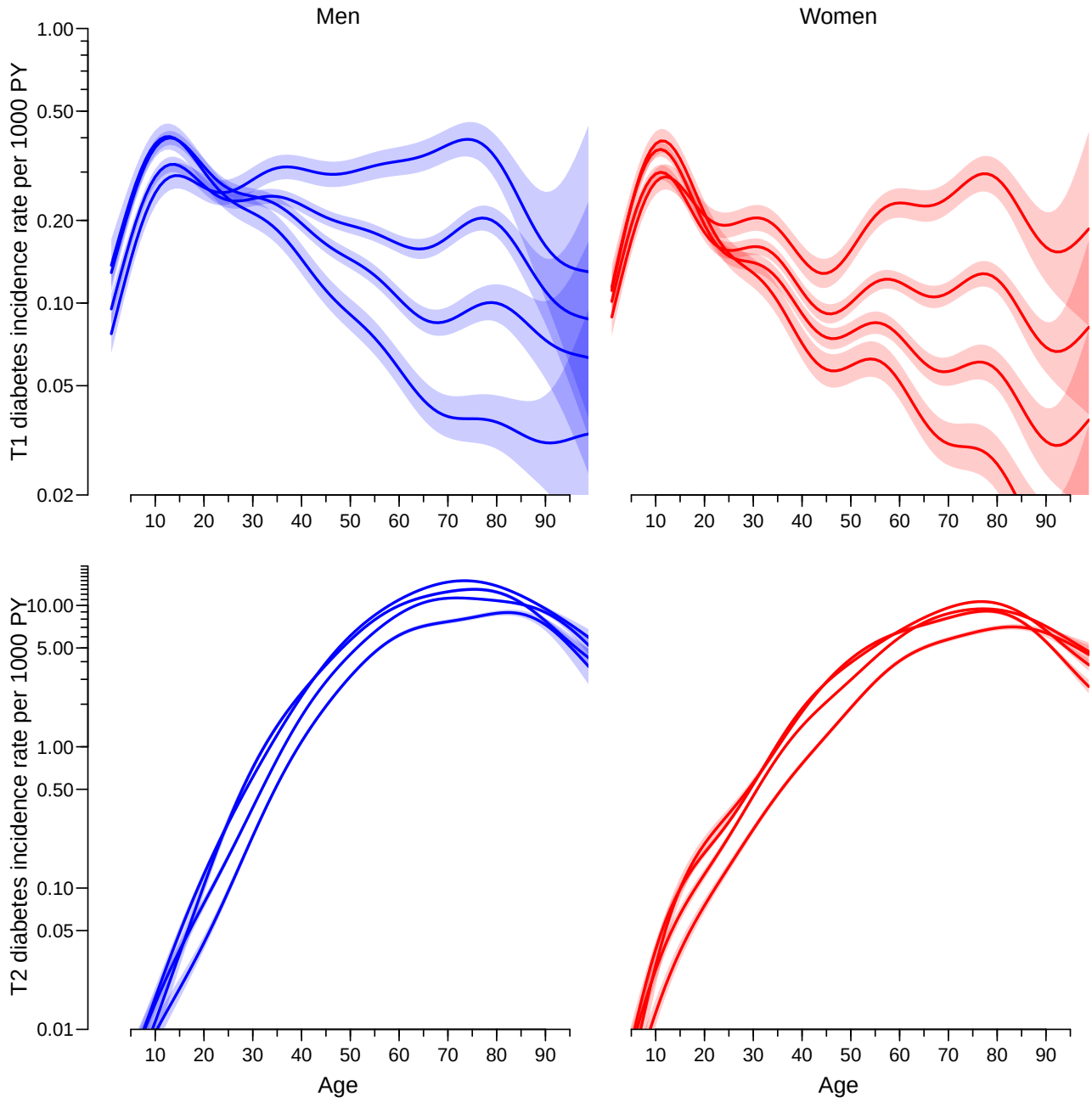


Figure 6.8: *Incidence rates of T1 and T2 diabetes in Denmark at different dates as predicted by* `gam` *models with age, period and cohort effects. Note the y-axes are different between T1 and T2, with different relative extent as well. Blue: men, red: women. Shaded areas are 95% confidence intervals.* `./graph/inc-age-gam`

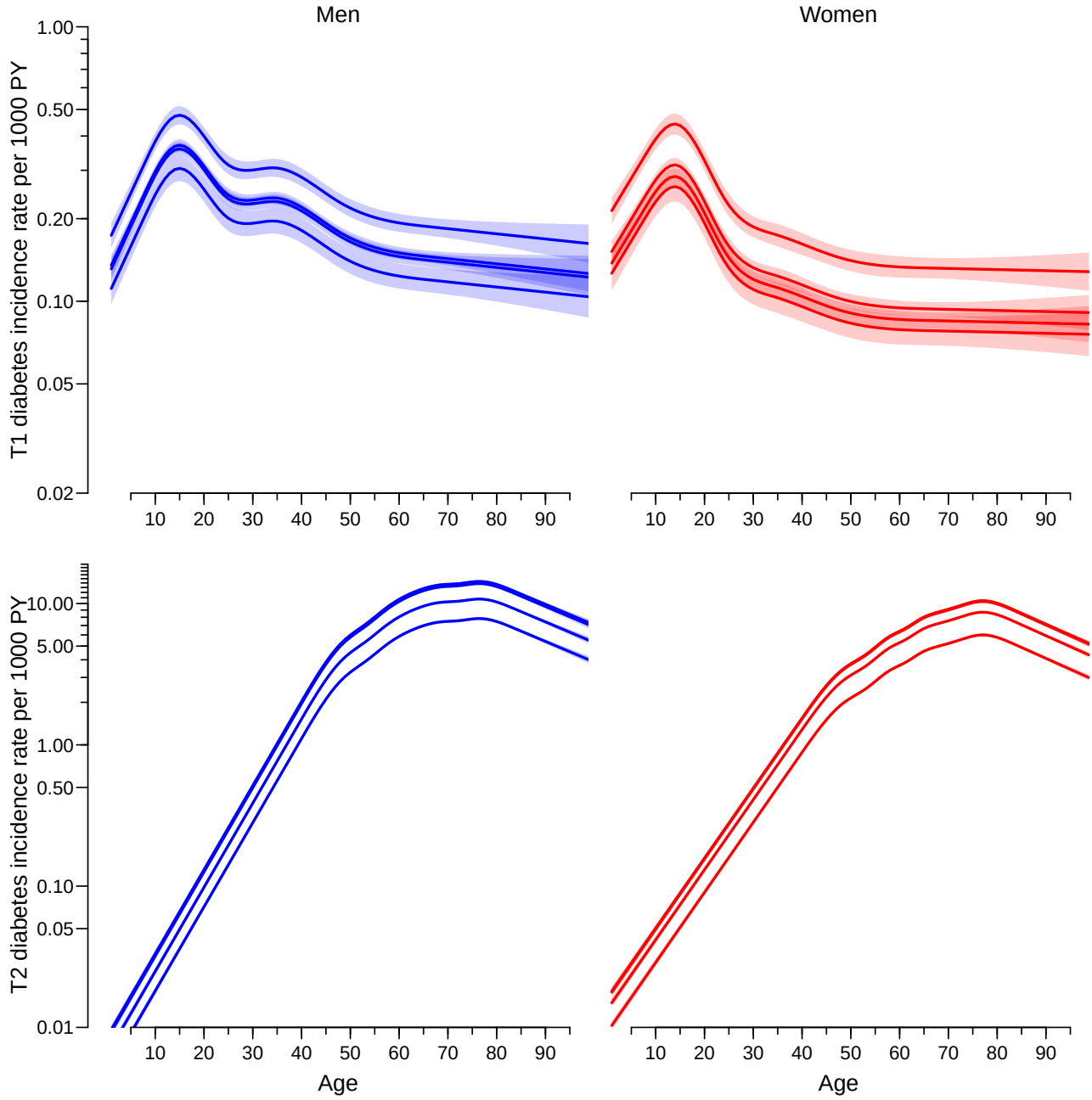


Figure 6.9: *Incidence rates of T1 and T2 diabetes in Denmark at different dates as predicted by an age-period model (hence the parallel curves). Note the y-axes are different between T1 and T2, with different relative extent as well. Blue: men, red: women. Shaded areas are 95% confidence intervals.* `./graph/inc-age-ap`

6.3.2 Rates by calendar time in different ages

Alternatively we could show the rates at different ages as a function of calendar time; we shall do this for ages 10, 20, 30,. . . ,70 for T1 and 30, 40,. . . ,90 for T2; note that we use the natural spline property of linearity to boldly predict rates beyond 2016-01-01. These predictions are slightly non-parallel for the models with cohort terms.

```
> pl.per <-
+ function( mod = "APC", nlt=NULL )
+    {
+ # utilities for letter-labeling
+ ulc <- function(){uu<-par("usr")
+        list(x=uu[1]*0.95+uu[2]*0.05, y=10^(uu[4]*0.98+uu[3]*0.02)) }
+ nxl <- function(){nlt<<-nlt+1
+        text(ulc(),letters[nlt],font=2,cex=1.5) }
+
+ par( mfrow=c(2,2), mar=c(3,2,0,0), oma=c(0,2,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ for( tp in dimnames(prArr)[["typ"]] )
+ for( sx in dimnames(prArr)[["sex"]] )
+    {
+    wh.a <- if( tp=="T1" ) paste(1:6*10) else paste(4:9*10)
+    incmat <- NULL
+    for( i in wh.a ) incmat <- cbind(incmat, prArr[i,,,tp,sx,mod] )
+    matshade( p.pt, incmat, plot=TRUE,
+              type="l", lwd=2, lty=1,
+              col=if(sx=="M") "blue" else "red",
+              xlab="", ylab="", xlim=c(1996,2019),
+              ylim=c(0.5,20)/(1 + (tp=="T1")*19),
+              yaxs="i", xaxt="n", yaxt="n", log="y" )
+    if( !is.null(nlt) ) nxl()
+    axis( side=1, at=seq(2000,2015,5 ) )
+    axis( side=1, at=1996:2017, tcl=-0.3, labels=NA )
+    axis( side=2)
+    axis( side=2, at=outer(c(1,1.5,2:9),-1:1-(tp=="T1"),function(x,y) x*10^y),
+          labels=NA, tcl=-0.3 )
+    if( sx=="M" ) mtext( paste( tp, "diabetes incidence rate per 1000 PY"),
+                         side=2, line=2.5, las=0 )
+    ecrv <- prArr[wh.a,"2017",1,tp,sx,mod]
+    text( 2018, lpos <- exp(seq(log(ecrv[1]),log(rev(ecrv)[1]),,length(ecrv))),
+                wh.a, adj=0, font=2 )
+    segments( 2017, ecrv, 2018, lpos )
+    }
+ mtext( "Date of follow-up", side=1, line=2, outer=TRUE )
+ mtext( c("Men","Women"), at=c(1,3)/4, side=3, line=0, outer=TRUE )
+ mtext( rep("Date of follow-up",2), at=c(1,3)/4, side=1, line=-1, outer=TRUE )
+    }
> pl.per("APC",nlt=0)
```

We can make the same plots with the `gam` models, but they are too wiggly to be really credible.

```
> pl.per("gam")
```

```
> pl.per("AP")
```

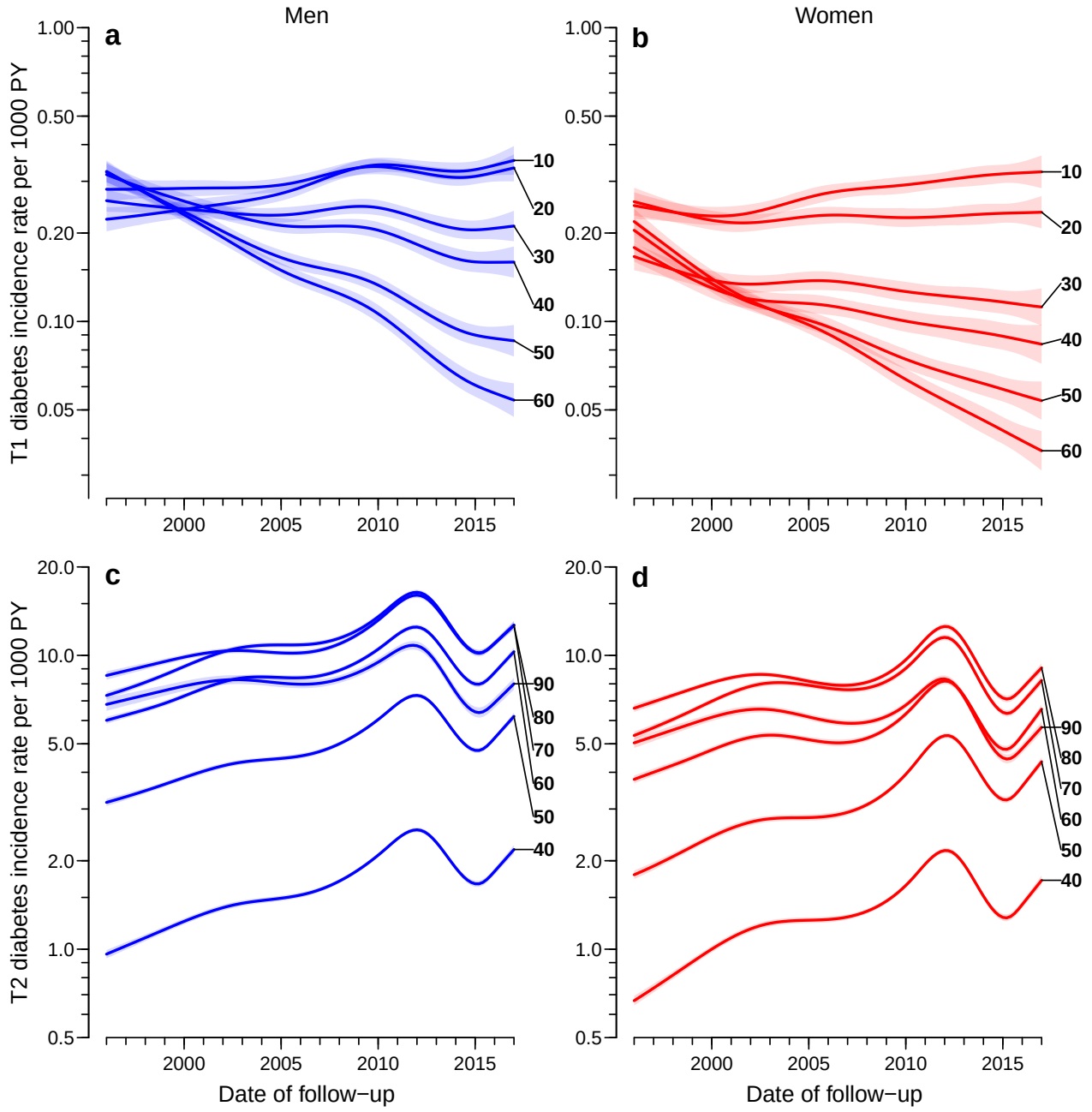


Figure 6.10: *Time trend in diabetes incidence rates in Denmark at different ages based on age-period-cohort models with spline effects with (6,5,5) knots for T1 and (10,5,10) knots for T2. Note the y-axes are different for T1 and T2. The* relative *extent of the axes are the same — a factor 30 from bottom to top. Blue: men, red: women. Shaded areas are 95% confidence intervals.* `./graph/inc-per-apc`

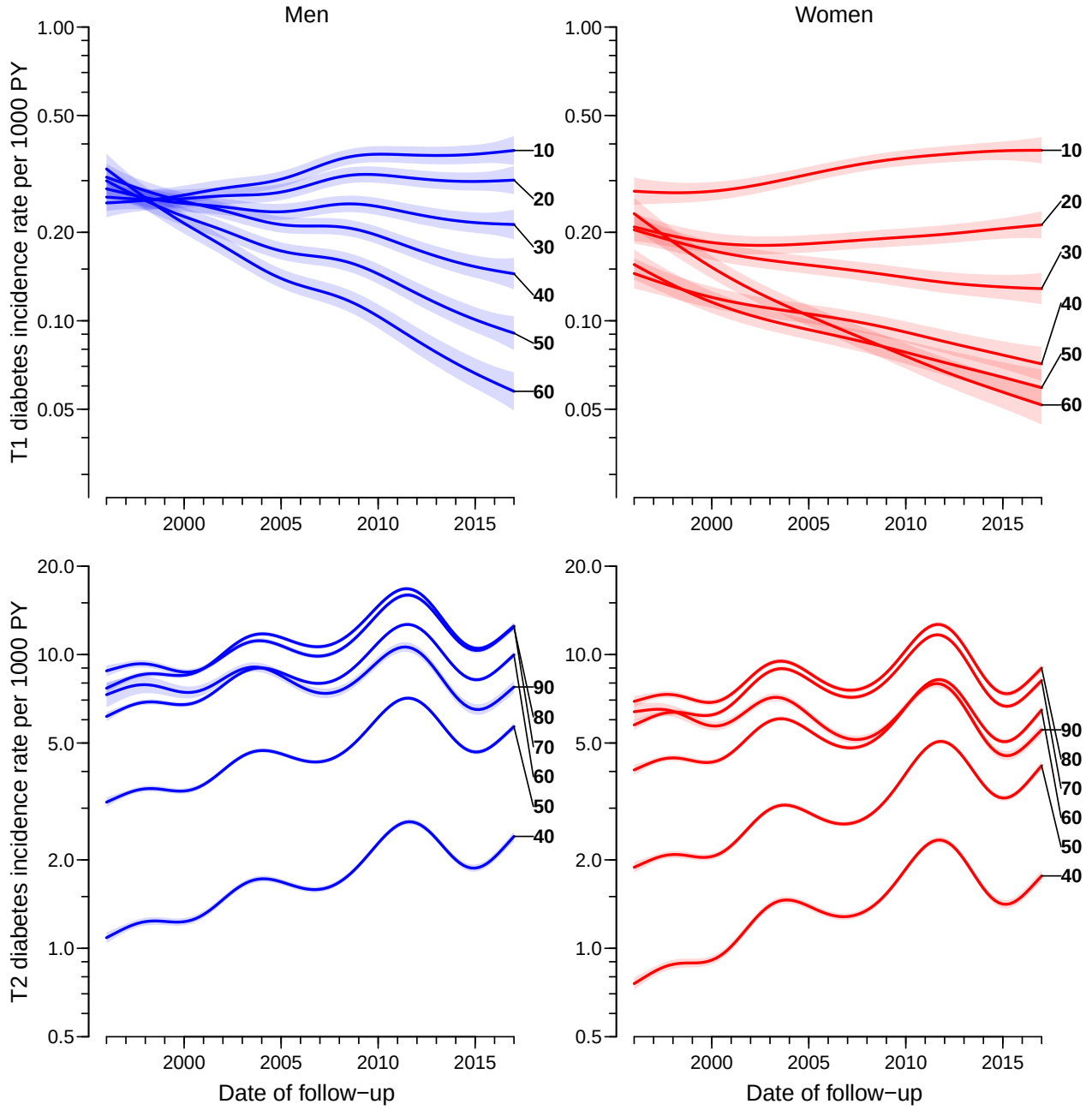


Figure 6.11: *Time trend in diabetes incidence rates in Denmark at different ages as predicted by* `gam` *models with age, period and cohort terms. Note that the ages are different for T1 and T2; but the extent of the axes are the same — a factor 30 from bottom to top. Blue: man, red: women. Shaded areas are 95% confidence intervals.* `./graph/inc-per-gam`

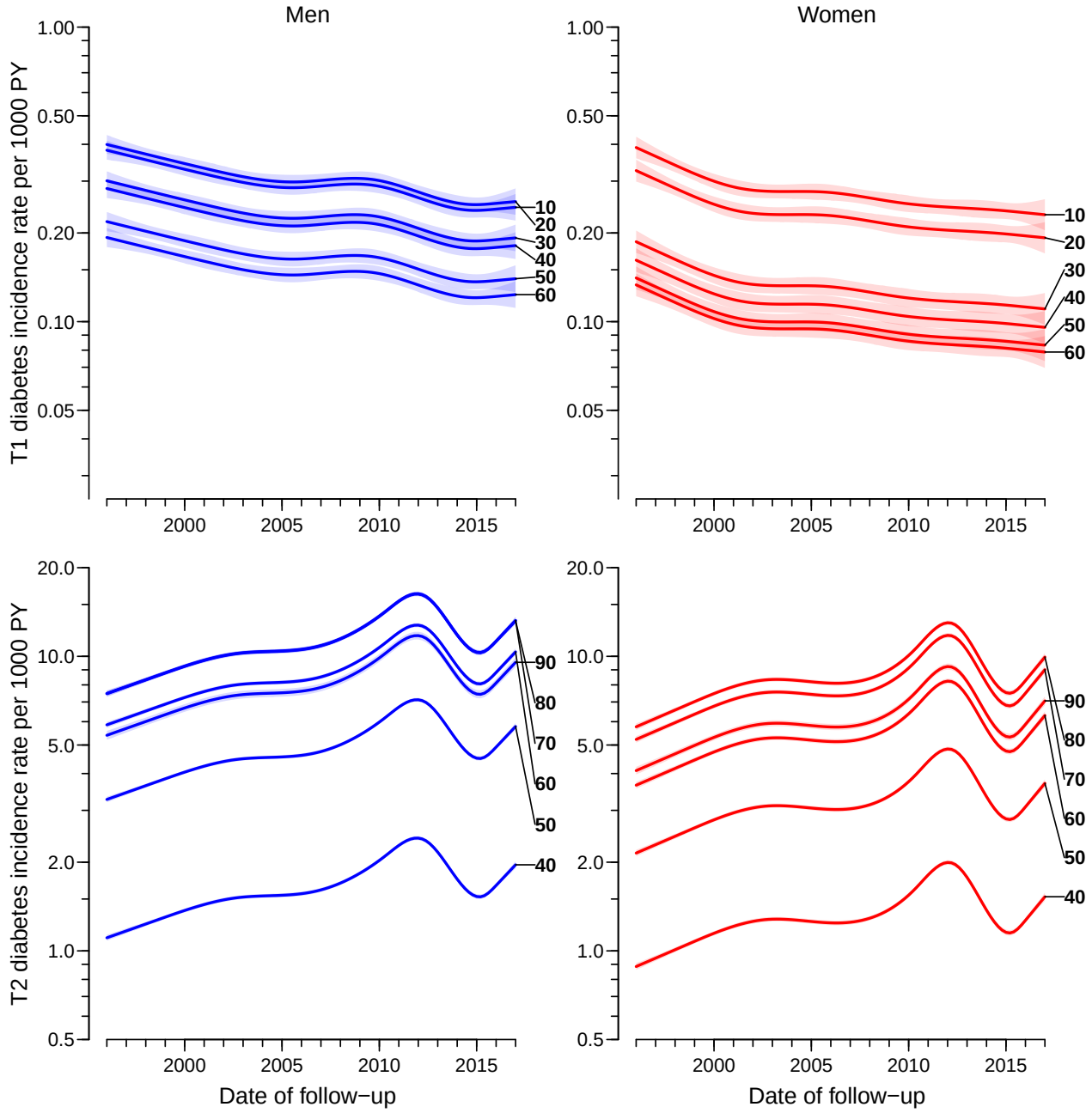


Figure 6.12: *Time trend in diabetes incidence rates in Denmark at different ages as predicted by an age-period model (hence the parallel lines). Note that the ages are different for T1 and T2; but the extent of the y-axes are the same — a factor 30 from bottom to top. Blue: man, red: women. Shaded areas are 95 confidence intervals.* `./graph/inc-per-ap`

6.4 Incidence summary

The incidence rates of T1D is about 0.3 per 1000 PY in ages 10 to 20, lower in older ages; at age 50 about third of this. There has been a substantial decrease in incidence rates of T1D in older persons over the last 20 years, for persons aged 50 some 8% year, wheres there has been an annual increase of 1–2% per year for persons under 20.

Incidence rates of T2D in ages over 60 is in the range of 6–10 per 1000 PY, at age 50 half of this, at age 40 a quarter of this. T2D incidence rates has been increasing overall about 3% per year, somewhat less in ages over 60, about 1% per year. The overall trend is somewhat misleading for T2D since a drop in incidence was seen during year 2012 through 2014 and in the subsequent increase after that.

6.4.1 Saving the fitted models

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

```
> save( M1apc, M1ac, M1ap, M1gam,
+       F1apc, F1ac, F1ap, F1gam,
+       M2apc, M2ac, M2ap, M2gam,
+       F2apc, F2ac, F2ap, F2gam,
+       file="../nydata/inc.Rda" )


> elapsed()
----------------------
2019-01-10 at 11:23:12
Time elapsed: 00:01:06
----------------------
```

. . . now input from `mort.tex`

Chapter 7

Mortality

First we load packages and the rate datasets:

```
> library( Epi )
> library( splines )
> library( mgcv )
> start()
---------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-01-10 17:55:43
---------------------------------------

> load( file="../data/inits.Rda" )
> 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.

We produce a brief overview of the new cases and the deaths among diabetes patients:

```
> fCtable( xtabs( cbind( DM = T1+T2,
+                        dDM = D*(state!="noDM"),
+                          0 = T1+T2-D*(state!="noDM") ) ~ P,
+                  data=rt ) )
              DM       dDM         0
P
1996     12,866     6,116     6,750
1997     12,633     6,306     6,327
1998     14,080     6,397     7,683
1999     14,723     6,841     7,882
2000     16,626     6,947     9,679
2001     14,591     7,054     7,537
2002     16,928     7,406     9,522
2003     18,608     7,747    10,861
2004     19,106     7,609    11,497
2005     17,796     7,902     9,894
2006     17,267     8,000     9,267
2007     18,637     8,158    10,479
2008     20,664     8,034    12,630
2009     20,459     8,716    11,743
2010     22,201     8,808    13,393
2011     25,385     8,839    16,546
2012     20,858     9,158    11,700
```

```
2013     16,372     9,431     6,941
2014     15,973     9,746     6,227
2015     17,849    10,079     7,770
2016     19,104    10,259     8,845
```

First we summarize the ages at death — the column `mod-80` gives the width of the age-range from the 10th to the 90th percentile of ages at death:

```
> dthA <- do.call( rbind,
+                  with( rt,
+                        tapply( rep(Ax,D),
+                                rep(state,D),
+                                quantile,
+                                c(1,5,9)/10 ) ) )
> dthA <- cbind( dthA, dthA[,3]-dthA[,1] )
> colnames( dthA )[4] <- "mid80"
> round( dthA, 1 )
      10%  50%  90% mid80
noDM 56.3 79.3 91.7  35.3
T1   48.3 71.7 87.3  39.0
T2   62.3 78.7 90.3  28.0
```

Not surprisingly the T2 deaths occur in a much narrower intervals, mainly because the persons at risk are much older.

7.1 Mortality overview

For convenience we make special data frames of T1 follow-up, one with and one without duration, but first an overview of deaths ex- resp. in-cluding the prevalent cases as of 1996-01-01:

```
> # Only persons diagnosed after 1996-01-01
> mL <- xtabs( D ~ floor(P) + state + sex, data=rtL )
> mL <- addmargins( mL, 2:3 )[,c(1:4,4),]
> dimnames( mL )[[2]][4] <- "DM"
> mL[,4,] <- mL[,2,] + mL[,3,]
> str( mL )
 'table' num [1:21, 1:5, 1:3] 26713 26055 25420 25041 24293 ...
 - attr(*, "dimnames")=List of 3
  ..$ floor(P): chr [1:21] "1996" "1997" "1998" "1999" ...
  ..$ state   : chr [1:5] "noDM" "T1" "T2" "DM" ...
  ..$ sex     : chr [1:3] "M" "W" "Sum"
> fCtable( mL[,  2:4 ,], w=6, col.vars=2:3 )
         state     T1                  T2                    DM
         sex        M      W    Sum      M      W    Sum      M      W    Sum
floor(P)
1996               14     12     26    255    222    477    269    234    503
1997               28     16     44    577    455  1,032    605    471  1,076
1998               50     30     80    860    715  1,575    910    745  1,655
1999               85     34    119  1,217    908  2,125  1,302    942  2,244
2000              101     58    159  1,435  1,180  2,615  1,536  1,238  2,774
2001               97     83    180  1,737  1,356  3,093  1,834  1,439  3,273
2002              142     70    212  1,929  1,616  3,545  2,071  1,686  3,757
2003              141    100    241  2,279  1,828  4,107  2,420  1,928  4,348
```

```
2004              157    102    259  2,349  1,968  4,317  2,506  2,070  4,576
2005              196    111    307  2,600  2,194  4,794  2,796  2,305  5,101
2006              189    129    318  2,736  2,335  5,071  2,925  2,464  5,389
2007              186    108    294  2,990  2,529  5,519  3,176  2,637  5,813
2008              206    128    334  3,083  2,536  5,619  3,289  2,664  5,953
2009              194    129    323  3,507  2,797  6,304  3,701  2,926  6,627
2010              199    125    324  3,664  2,970  6,634  3,863  3,095  6,958
2011              166    107    273  3,831  2,999  6,830  3,997  3,106  7,103
2012              151    105    256  4,159  3,138  7,297  4,310  3,243  7,553
2013              147     84    231  4,336  3,341  7,677  4,483  3,425  7,908
2014              114     95    209  4,613  3,544  8,157  4,727  3,639  8,366
2015              123     81    204  4,796  3,779  8,575  4,919  3,860  8,779
2016              133     76    209  4,988  3,874  8,862  5,121  3,950  9,071
```

```
> fCtable( mL[,-(2:4),], w=6, col.vars=2:3 )
```

```
         state   noDM                   Sum
         sex        M      W    Sum      M      W    Sum
floor(P)
1996           26,713 27,126 53,839 26,982 27,360 54,342
1997           26,055 26,965 53,020 26,660 27,436 54,096
1998           25,420 26,129 51,549 26,330 26,874 53,204
1999           25,041 26,930 51,971 26,343 27,872 54,215
2000           24,293 25,913 50,206 25,829 27,151 52,980
2001           24,485 26,249 50,734 26,319 27,688 54,007
2002           24,099 26,375 50,474 26,170 28,061 54,231
2003           23,846 25,434 49,280 26,266 27,362 53,628
2004           23,010 24,266 47,276 25,516 26,336 51,852
2005           22,324 24,042 46,366 25,120 26,347 51,467
2006           22,373 23,749 46,122 25,298 26,213 51,511
2007           22,325 24,182 46,507 25,501 26,819 52,320
2008           21,872 23,243 45,115 25,161 25,907 51,068
2009           21,650 23,358 45,008 25,351 26,284 51,635
2010           21,390 22,698 44,088 25,253 25,793 51,046
2011           20,460 21,834 42,294 24,457 24,940 49,397
2012           20,044 21,535 41,579 24,354 24,778 49,132
2013           20,052 21,131 41,183 24,535 24,556 49,091
2014           19,501 20,443 39,944 24,228 24,082 48,310
2015           20,045 20,902 40,947 24,964 24,762 49,726
2016           19,824 20,819 40,643 24,945 24,769 49,714
```

```
> # All persons
> mA <- xtabs( D ~ floor(P) + state + sex, data=rt  )
> mA <- addmargins( mA, 2:3 )[,c(1:4,4),]
> dimnames( mA )[[2]][4] <- "DM"
> mA[,4,] <- mA[,2,] + mA[,3,]
> str( mA )
```

```
 'table' num [1:21, 1:5, 1:3] 26713 26055 25420 25041 24293 ...
 - attr(*, "dimnames")=List of 3
  ..$ floor(P): chr [1:21] "1996" "1997" "1998" "1999" ...
  ..$ state   : chr [1:5] "noDM" "T1" "T2" "DM" ...
  ..$ sex     : chr [1:3] "M" "W" "Sum"
```

```
> fCtable( mA[,  2:4 ,], w=6, col.vars=2:3 )
```

```
         state     T1                   T2                   DM
         sex        M      W    Sum      M      W    Sum      M      W    Sum
floor(P)
1996              363    334    697  2,798  2,621  5,419  3,161  2,955  6,116
1997              442    311    753  2,819  2,734  5,553  3,261  3,045  6,306
```

```
1998              420    340    760  2,928  2,709  5,637  3,348  3,049  6,397
1999              453    331    784  3,193  2,864  6,057  3,646  3,195  6,841
2000              453    361    814  3,168  2,965  6,133  3,621  3,326  6,947
2001              433    357    790  3,276  2,988  6,264  3,709  3,345  7,054
2002              513    323    836  3,453  3,117  6,570  3,966  3,440  7,406
2003              473    381    854  3,667  3,226  6,893  4,140  3,607  7,747
2004              466    319    785  3,655  3,169  6,824  4,121  3,488  7,609
2005              487    329    816  3,724  3,362  7,086  4,211  3,691  7,902
2006              450    337    787  3,832  3,381  7,213  4,282  3,718  8,000
2007              425    278    703  3,911  3,544  7,455  4,336  3,822  8,158
2008              382    261    643  3,984  3,407  7,391  4,366  3,668  8,034
2009              357    256    613  4,464  3,639  8,103  4,821  3,895  8,716
2010              348    235    583  4,452  3,773  8,225  4,800  4,008  8,808
2011              311    200    511  4,614  3,714  8,328  4,925  3,914  8,839
2012              258    170    428  4,926  3,804  8,730  5,184  3,974  9,158
2013              240    143    383  5,054  3,994  9,048  5,294  4,137  9,431
2014              180    135    315  5,327  4,104  9,431  5,507  4,239  9,746
2015              185    118    303  5,431  4,345  9,776  5,616  4,463 10,079
2016              174    124    298  5,598  4,363  9,961  5,772  4,487 10,259
```

```
> fCtable( mA[,-(2:4),], w=6, col.vars=2:3 )
```

```
         state   noDM                    Sum
         sex        M      W    Sum       M      W    Sum
floor(P)
1996            26,713 27,126 53,839 29,874 30,081 59,955
1997            26,055 26,965 53,020 29,316 30,010 59,326
1998            25,420 26,129 51,549 28,768 29,178 57,946
1999            25,041 26,930 51,971 28,687 30,125 58,812
2000            24,293 25,913 50,206 27,914 29,239 57,153
2001            24,485 26,249 50,734 28,194 29,594 57,788
2002            24,099 26,375 50,474 28,065 29,815 57,880
2003            23,846 25,434 49,280 27,986 29,041 57,027
2004            23,010 24,266 47,276 27,131 27,754 54,885
2005            22,324 24,042 46,366 26,535 27,733 54,268
2006            22,373 23,749 46,122 26,655 27,467 54,122
2007            22,325 24,182 46,507 26,661 28,004 54,665
2008            21,872 23,243 45,115 26,238 26,911 53,149
2009            21,650 23,358 45,008 26,471 27,253 53,724
2010            21,390 22,698 44,088 26,190 26,706 52,896
2011            20,460 21,834 42,294 25,385 25,748 51,133
2012            20,044 21,535 41,579 25,228 25,509 50,737
2013            20,052 21,131 41,183 25,346 25,268 50,614
2014            19,501 20,443 39,944 25,008 24,682 49,690
2015            20,045 20,902 40,947 25,661 25,365 51,026
2016            19,824 20,819 40,643 25,596 25,306 50,902
```

The analysis datasets for T1D mortality:

```
> rt1 <- subset( rtL, state=="T1" )
> rx1 <- subset( rt , state=="T1" )
> str( rt1 )
'data.frame':        91349 obs. of  17 variables:
 $ sex  : Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ state: Factor w/ 3 levels "noDM","T1","T2": 2 2 2 2 2 2 2 2 2 2 ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ P    : num  1996 1997 1997 1997 1998 ...
 $ C    : num  1995 1996 1997 1997 1997 ...
 $ dur  : num  0.1 0.1 0.1 0.35 0.35 0.75 0.1 0.35 0.75 0.1 ...
```

```
 $ Dcvd : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dcan : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dres : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Doth : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y    : num  2.00e-04 1.79e-04 2.00e-04 2.61e-04 3.87e-05 ...
 $ T1   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Ax   : num  0.667 0.667 0.333 0.333 0.667 ...
 $ Px   : num  1996 1997 1998 1998 1998 ...
 $ Cx   : num  1996 1997 1997 1997 1998 ...
```

```
> str( rx1 )
```

```
'data.frame':        196511 obs. of  17 variables:
 $ sex  : Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ state: Factor w/ 3 levels "noDM","T1","T2": 2 2 2 2 2 2 2 2 2 2 ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ P    : num  1996 1996 1997 1997 1997 ...
 $ C    : num  1995 1995 1996 1997 1997 ...
 $ dur  : num  0.1 0.75 0.1 0.1 0.35 0.35 0.75 0.1 0.35 0.75 ...
 $ Dcvd : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dcan : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dres : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Doth : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y    : num  0.0002 0.000283 0.000179 0.0002 0.000261 ...
 $ T1   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Ax   : num  0.667 0.667 0.667 0.333 0.333 ...
 $ Px   : num  1996 1996 1997 1998 1998 ...
 $ Cx   : num  1996 1996 1997 1997 1997 ...
```

7.2 Type 1 diabetes patient mortality

In these 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 for men and women:

```
> ( 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%
41.33333 54.66667 62.33333 67.66667 73.33333 78.66667 83.66667 90.33333
```

```
> ( d.kn <- c(0,with( rt1, quantile( rep(   dur,D), (1:3-0.0)/4 ) )) )
    25% 50% 75%
0.0 1.5 4.5 8.5
```

```
> ( e.kn <-     with( rt1, quantile( rep(Ax-dur,D), (1:5-0.5)/5 ) )  )
     10%      30%      50%      70%      90%
41.16667 55.83333 64.83333 73.83333 82.16667
```

```
> ( p.kn <-     with( rt1, quantile( rep(   Px ,D), (1:5-0.5)/5 ) )  )
     10%      30%      50%      70%      90%
2001.333 2005.333 2008.333 2011.333 2014.667
```

```
> ( 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%
 1916.667  1923.222  1927.667  1932.667  1937.667  1942.667  1947.667  1954.778  1969.333
```

With these knot-vectors in place we set up separate mortality models for men and women, simple age-period models, as well as more elaborate models taking age at diagnosis and duration of diabetes into account.

```
> 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=="W" ) )
> mtx <- update( mt0, data = subset( rx1, sex=="M" ) )
> ftx <- update( mt0, data = subset( rx1, sex=="W" ) )
> 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=="W" ) )
> rt1$Ae <- rt1$Ax - rt1$dur
> ml1 <- update( mt1, . ~ . - Ns( Px, knots=p.kn ) + Px )
> fl1 <- update( ft1, . ~ . - Ns( Px, knots=p.kn ) + Px )
> round( ( rbind( "M"=ci.exp( ml1, subset="Px" )[1,],
+                 "W"=ci.exp( fl1, subset="Px" )[1,] ) - 1 )*100, 1 )
  exp(Est.) 2.5% 97.5%
M      -0.5 -1.3   0.3
W      -0.2 -1.2   0.9
```

So we see there is an overall *in*crease in mortality among T1D patients.

As a final check we fit a cohort-effect to the residuals from the age, date duration and period model, to check if there is a residual cohort-effect:

```
> prAP <- predict(mt1,type="response")
> mr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(prAP),
+             family = poisson,
+               data = subset(rt1,sex=="M") )
> prAP <- predict(ft1,type="response")
> fr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(prAP),
+             family = poisson,
+               data = subset(rt1,sex=="W") )
```

We then set up data frames for predicting the mortality rates for men and women as of 2015-01-01, for different combinations of age at diagnosis (entry `e.pr`), date of diagnosis 2000 and a sequence of durations:

```
> e.pr <- seq(15,65,5)
> d.pr <- c(NA,seq(0,20,0.1))
> nd0 <- ndx <- data.frame( Ax = seq( 5,90,0.2), Px=2015, Y=1 )
> nd  <- data.frame( expand.grid( dur=d.pr, Ae=e.pr ) )
> nd <- transform( nd, Ax = Ae+dur,
+                      Px = 2015,
+                       Y = 1 )
> head( nd )
```

```
  dur Ae   Ax   Px Y
1  NA 15   NA 2015 1
2 0.0 15 15.0 2015 1
3 0.1 15 15.1 2015 1
4 0.2 15 15.2 2015 1
5 0.3 15 15.3 2015 1
6 0.4 15 15.4 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` (not for matshade till `Epi_2.32`, though):

```
> t0pr <- cbind( ci.pred( mt0, nd0 ), ci.pred( ft0, nd0 ) )
> t0pr <- cbind( t0pr, ci.ratio(t0pr[,1:3],t0pr[,1:3+3]) )
> txpr <- cbind( ci.pred( mtx, ndx ), ci.pred( ftx, ndx ) )
> txpr <- cbind( txpr, ci.ratio(txpr[,1:3],txpr[,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)",
+       xlim=c(10,85), ylim=c(0.2,200), yaxs="i" )
> abline( v=e.pr, col=gray(0.8) )
> abline( h=1 )
> axis( side=1, at=seq(10,85,10), labels=NA, tcl=-0.4 )
> axis( side=1, at=seq(10,85, 5), labels=NA, tcl=-0.2 )
> axis( side=2, at=outer(2:10,-1:2,function(x,y) x*10^y), labels=NA, tcl=-0.3 )
> for( e in e.pr[] )
+ matshade( nd$Ax[nd$Ae==e], t1pr[nd$Ae==e,],
+           type="l", lwd=2, lty=1,
+           col=c("blue","red",gray(0.3)), )
> matshade( nd0$Ax, t0pr,
+           type="l", lwd=3, lty="11", lend="butt",
+           col=c("blue","red",gray(0.5)) )
> matshade( ndx$Ax, txpr,
+           type="l", lwd=1, lty=1,
+           col=c("blue","red",gray(0.5)) )
> 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 )
```

From figure 8.1 it is pretty obvious that longer diabetes duration (or younger age at diagnosis) is associated with higher mortality at a given age.

Here are a few figures for the talk:

```
> plex <-
+ function( e, cols, s )
+ {
+ pdf(paste("./graph/slides-",e,s,".pdf",sep=""),width=8,height=5)
+ par( mar=c(3,4,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ plot( NA,
+       log="y", xlab="Age at follow-up",
+                ylab="",
+       xlim=c(10,85), ylim=c(0.2,200), yaxs="i" )
+ mtext( "T1D: mortality per 1000 PY (2015)", side=2, line=3, las=0 )
+ abline( v=e, col=gray(0.8) )
```

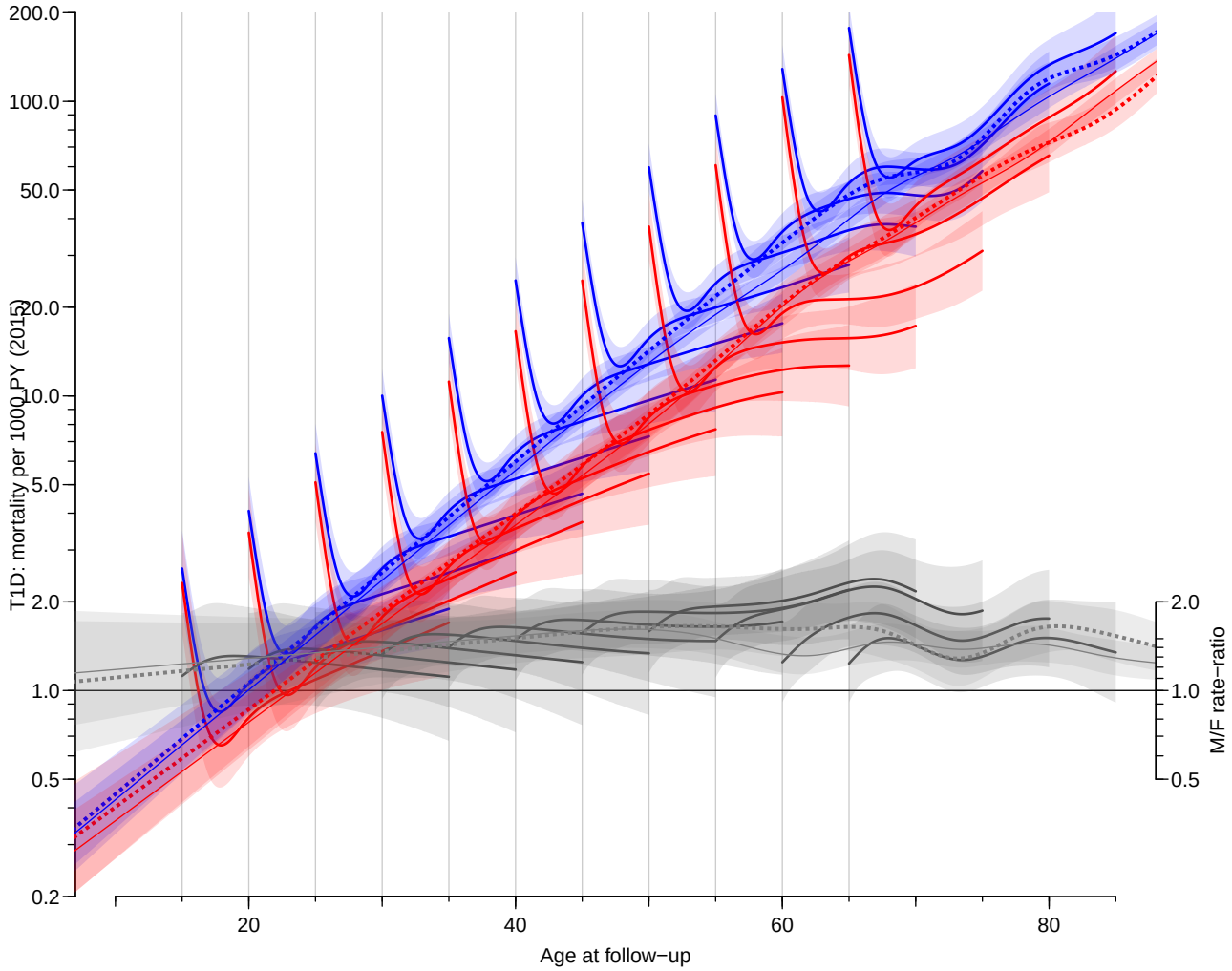


Figure 7.1: *Rates of T1D mortality by age and duration of diabetes, at 2015-01-01. Each full curve corresponds to a fixed age at diagnosis, starting at the age at diagnosis. Dotted curves are from the model where age at diagnosis and duration is ignored, the thin full lines are from including prevalent cases as of 1996-01-01 in the analysis ignoring diabetes duration. Red curves are women, blue men, black curves are the M/W rate ratios. The shaded areas indicate 95% confidence bands.* `./graph/mort-amort-t1`

```
+ # abline( h=1 )
+ axis( side=1, at=seq(10,85,10), labels=NA, tcl=-0.4 )
+ axis( side=1, at=seq(10,85, 5), labels=NA, tcl=-0.2 )
+ axis( side=2, at=outer(2:10,-1:2,function(x,y) x*10^y), labels=NA, tcl=-0.3 )
+ matshade( nd$Ax[nd$Ae==e], t1pr[nd$Ae==e,cols],
+           type="l", lwd=2, lty=1,
+           col=c("blue","red") )
+ axis( side=1, at=e+0:4*5, labels=0:4*5, line=-5 )
+ axis( side=1, at=e+0:20, line=-5, labels=NA, tcl=-0.3 )
+ mtext( side=1, at=e+12, "Duration of diabetes", line=-3 )
+ text( e+1, 100, paste("Diagnosed age", e ), adj=0 )
+ dev.off()
+ }
> plex( 15, 1:3, "M" )
```

```
pdf
  2
> plex( 20, 1:3, "M" )
pdf
  2
> plex( 25, 1:3, "M" )
pdf
  2
> plex( 30, 1:3, "M" )
pdf
  2
> plex( 35, 1:3, "M" )
pdf
  2
> plex( 40, 1:3, "M" )
pdf
  2
> plex( 15, 1:6, "MW" )
pdf
  2
> plex( 20, 1:6, "MW" )
pdf
  2
> plex( 25, 1:6, "MW" )
pdf
  2
> plex( 30, 1:6, "MW" )
pdf
  2
> plex( 35, 1:6, "MW" )
pdf
  2
> plex( 40, 1:6, "MW" )
pdf
  2
```

Then we plot the period-RR, extracting the values from the model object, and also showing the cohort RRs as a check of the model fit:

```
> par( mfrow=c(1,2), mar=c(3,3,1,3), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pr.P <- seq(1996,2017,0.2)
> ct.lst <- list( data.frame(Px=pr.P),
+                 data.frame(Px=2015) )
> xv <- c("Ax","dur")
> t1rr <- cbind( ci.exp( mt1, ct.lst, xvar=xv),
+                ci.exp( ft1, ct.lst, xvar=xv) )
> plot( NA, log="y", xlim=c(1996,2017), ylim=c(0.5,2),
+       xlab="Date of follow-up",
+       ylab="T1D: mortality RR relative to 2015" )
```

```
> abline( h=1 )
> axis( side=1, at=1996:2017, labels=NA, tcl=-0.3 )
> axis( side=2, at=5:15/10, labels=NA, tcl=-0.3 )
> matshade( pr.P, t1rr, lwd=2, lty=1, col=c("blue","red") )
> pr.C <- 1910:1990
> nd <- data.frame( Px=2015, Ax=2015-pr.C, prAP=1 )
> t1cr <- cbind( ci.pred( mr1, nd ),
+                ci.pred( fr1, nd ) )
> 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 )
> axis( side=1, at=seq(1910,1990,10), labels=NA, tcl=-0.3 )
> axis( side=2, at=5:15/10, labels=NA, tcl=-0.3 )
> matshade( pr.C, t1cr, lwd=2, lty=1, col=c("blue","red") )
```

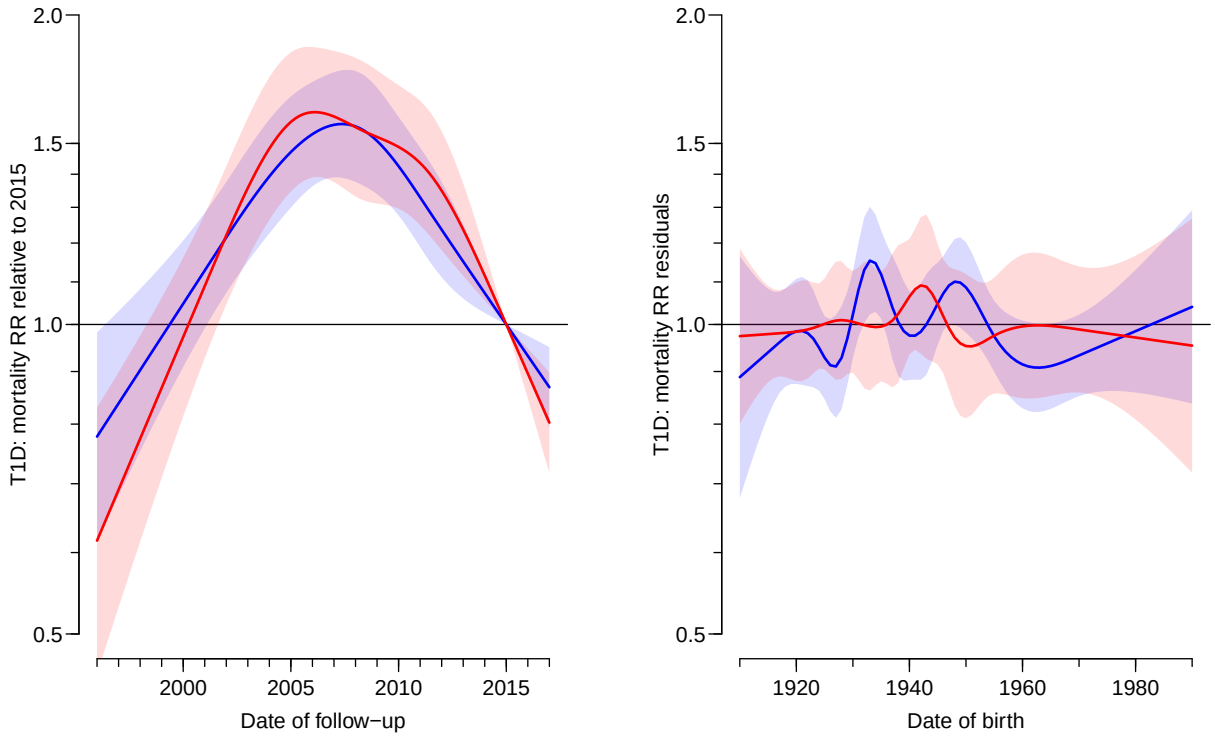


Figure 7.2: *Left panel: Rate-ratio of T1D mortality relative to 2015-01-01. Right panel: Residuals modeled by date of birth (cohort). Shaded areas indicate 95% confidence intervals; red curves are women, blue men.* `./graph/mort-prr-crr-t1`

From figure 8.2 we se a remarkable increase in mortality from 1996 to 2010, and then an abrupt decline. The model is a model with effect of age at diagnosis, age at follow-up and duration. This may be an indication of misclassification of T1D as T2D in the period before the start of the DADD (in 2005) leaving only the survivors to be classified as T1D, conveying a smaller mortality in the earlier period.

The cohort residuals are shown as a check of model-fit, and show no indication of any systematic effect of date of birth..

7.3 Mortality for T1, T2 and all DM

Similar analyses as those above will be done for T2 patients as well as for T1 and T2 combined. Hence it will be suitable to automate it all in a single function that returns the estimated rates and RRs in a list, which then can be used to produce combined plots

```
> mort.res <-
+ function( rt1, rx1 )
+     {
+ # 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), (1:3-0.0)/4 ) ))
+ 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,
+               data = subset( rt1, sex=="M" ) )
+ ft0 <- update( mt0, data = subset( rt1, sex=="W" ) )
+ 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=="W" ) )
+ mtx <- update( mt0, data = subset( rx1, sex=="M" ) )
+ ftx <- update( mt0, data = subset( rx1, sex=="W" ) )
+ 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","W")
+ 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["W","P(lin)"] <- anova( ft1, fl1, test="Chisq" )[2,5]
+ test.tr["M","P(sl)"] <- ci.lin( ml1, subset="Px" )[,"P"]
+ test.tr["W","P(sl)"] <- ci.lin( fl1, subset="Px" )[,"P"]
+ test.tr["M",c(3,5,4)] <- (ci.exp( ml1, subset="Px" ) - 1)*100
+ test.tr["W",c(3,5,4)] <- (ci.exp( fl1, subset="Px" ) - 1)*100
+
+ # Residuals by cohort
+ prAP <- predict(mt1,type="response")
+ mr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(prAP),
+             family = poisson,
+               data = subset(rt1,sex=="M") )
+ prAP <- predict(ft1,type="response")
+ fr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(prAP),
+             family = poisson,
```

```
+               data = subset(rt1,sex=="W") )
+
+ # Prediction data frames
+ e.pr <- seq(10,75,5)
+ d.pr <- c( NA, seq(0.1,20,0.1) )
+ 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 and M/F RRs
+ t0pr <- cbind( ci.pred( mt0, nd0 ), ci.pred( ft0, nd0 ) )
+ t0pr <- cbind( t0pr, ci.ratio(t0pr[,1:3],t0pr[,1:3+3]) )
+ txpr <- cbind( ci.pred( mtx, nd0 ), ci.pred( ftx, nd0 ) )
+ txpr <- cbind( txpr, ci.ratio(txpr[,1:3],txpr[,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
+ ct.lst <- list( data.frame(Px=seq(1996,2017,0.2)),
+                 data.frame(Px=2015) )
+ xv <- c("Ax","dur")
+ t1rr <- cbind( ci.exp( mt1, ct.lst, xvar=xv),
+                ci.exp( ft1, ct.lst, xvar=xv) )
+
+ # Cohort residuals
+ ndc <- data.frame( Px=2015, Ax=2015-(1910:1990), prAP=1 )
+ t1cr <- cbind( ci.pred( mr1, ndc ),
+                ci.pred( fr1, ndc ) )
+
+ # return results
+ list( tr = test.tr,
+       A0 = data.frame(nd0[,"Ax"],t0pr,txpr),
+       Ad = data.frame(nd[,c("Ae","dur","Ax")],t1pr),
+       Pr = data.frame(P=pr.P,t1rr),
+       Cr = data.frame(C=pr.C,t1cr) )
+ }
> # Now use the function to generate results for T1, T2 and combined DM
> rT1 <- mort.res( subset( rtL, state=="T1" ),
+                  subset( rt , state=="T1" ) )
> rT2 <- mort.res( subset( rtL, state=="T2" ),
+                  subset( rt , state=="T2" ) )
> rDM <- mort.res( subset( rtL, state %in% c("T1","T2") ),
+                  subset( rt , state %in% c("T1","T2") ) )
> save( rT1, rT2, rDM, file="../data/mort-res.Rda" )
```

An overview of the annual changes in mortality:

```
> load( file="../data/mort-res.Rda" )
> fCp( rT1$tr, d=3, z=NULL )
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.000     0.234    -0.507     0.329    -1.337
W     0.000     0.748    -0.168     0.861    -1.186

> fCp( rT2$tr, d=3, z=NULL )
```

```
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.000     0.000    -3.302    -3.127    -3.477
W     0.000     0.000    -2.529    -2.330    -2.728
> fCp( rDM$tr, d=3, z=NULL )
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.000     0.000    -3.387    -3.216    -3.557
W     0.000     0.000    -2.570    -2.375    -2.765
```

7.4 Plotting mortality rates

We will make two sets of plots with 2 rows indexed by T1 / T2; one set with mortality, period RR and cohort residuals aimed for the ESM and one with only mortality rates and period RR aimed at the paper itself.

```
> pltp <-
+ function( rT1, coh=FALSE, e.pr=seq(10,70,10),
+           lab="", rlim=c(0.2,200), rrpos=1, nlt=NULL, rr=FALSE, article=TRUE )
+     {
+ # utilities for letter-labeling
+ ulc <- function(){uu<-par("usr")
+        list(x=uu[1]*0.93+uu[2]*0.07, y=10^(uu[4]*0.97+uu[3]*0.03)) }
+ nxl <- function(){nlt<<-nlt+1
+        text(ulc(),paste(letters[nlt],lab,sep="   "),font=1,cex=1.3,adj=0) }
+
+ rdf <- subset( rT1$Ad, Ae %in% e.pr )
+ art <- rT1$A0
+ prt <- rT1$Pr
+ crt <- rT1$Cr
+
+ # Adjust the RRs
+ rdf[,10:12] <- rdf[,10:12] * rrpos
+ art[,c(8:10,17:19)] <- art[,c(8:10,17:19)] * rrpos
+
+ # nice plotting frame for rates
+ plot( NA, log="y",
+       xlim=c(10,85), ylim=rlim, yaxs="i", xlab="", ylab="" )
+ abline( v=e.pr, col=gray(0.8) )
+ abline( h=rrpos )
+ if( rr ) abline( h=1 )
+ # axis grooming
+ axis( side=1, at=seq(10,85,10), labels=NA, tcl=-0.4 )
+ axis( side=1, at=seq(10,85, 5), labels=NA, tcl=-0.2 )
+ axis( side=2, at=outer(2:10,-1:2,function(x,y) x*10^y), labels=NA, tcl=-0.3 )
+ axis( side=4, at=c(5,10,20)/10*rrpos, labels=c(5,10,20)/10 )
+ axis( side=4, at=c(5:9,15,20)/10*rrpos, labels=NA, tcl=-0.3 )
+ mtext( "M/W\nRR", side=4, at=rrpos*3, line=0 )
+ # rates plotted
+ for( e in e.pr )
+ matshade( rdf$Ax[rdf$Ae==e], rdf[rdf$Ae==e,-(1:3)],
+           type="l", lwd=2, lty=1,
+           col=c("blue","red",gray(0.3)), )
+ if( article ) {
+ # matshade( art[,1], art[,2:10],
+ matlines( art[,1], art[,c(2,5,8)],
```

```
+           type="l", lwd=3, lty="11", lend="butt",
+           col=c("blue","red",gray(0.5)) )
+ # matshade( art[,1], art[,11:19],
+ matlines( art[,1], art[,c(11,14,17)],
+           type="l", lwd=1, lty=1, lend="butt",
+           col=c("blue","red",gray(0.5)) )
+           }
+ if( !is.null(nlt) ) nxl()
+
+ plot( NA, log="y", xlim=c(1996,2017), ylim=c(0.4,2.5),
+       xlab="", ylab="" )
+ abline( h=1 )
+ axis( side=1, at=1996:2017, labels=NA, tcl=-0.3 )
+ axis( side=2, at=4:15/10, labels=NA, tcl=-0.3 )
+ matshade( prt$P, prt[,-1], lwd=2, lty=1, col=c("blue","red") )
+ if( !is.null(nlt) ) nxl()
+
+ if( coh ){
+ plot( NA, log="y", xlim=c(1910,1990), ylim=c(0.4,2.5),
+       xlab="", ylab="" )
+ abline( h=1 )
+ axis( side=1, at=seq(1910,1990,10), labels=NA, tcl=-0.3 )
+ axis( side=2, at=4:15/10, labels=NA, tcl=-0.3 )
+ matshade( crt$C, crt[,-1], lwd=2, lty=1, col=c("blue","red") )
+ if( !is.null(nlt) ) nxl()
+          }
+     }
```

With this function defined we plot a 2 by 3 layout aimed at the ESM:

```
> layout( matrix(1:6,2,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rT1, coh=TRUE, nlt=0, e.pr=1:5*10+5)
> pltp( rT2, coh=TRUE, nlt=3, e.pr=3:7*10)
> mtext( c("Mortality rate per 1000 PY",
+          "RR relative to 2015",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE, cex=0.67 )
> 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(1:2,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(1,1,1,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rT1, coh=FALSE, lab="T1D", nlt=0, e.pr=1:5*10+5,
+       rlim=c(0.1,200), rrpos=0.2, article=FALSE )
> mtext( c("Mortality rate per 1000 PY",
+          "Rate ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

```
> layout( matrix(1:2,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(1,1,1,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rT2, coh=FALSE, lab="T2D", nlt=0, e.pr=3:7*10,
+       rlim=c(0.1,200), rrpos=0.2, article=FALSE )
> mtext( c("Mortality rate per 1000 PY",
+          "Rate ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

For the paper we leave out the period effects and cohort residuals:

```
> pltpa <-
+ function( rT1, e.pr=seq(10,70,10), tf=1,
+           lab="", rlim=c(0.2,200), nlt=NULL, rr=FALSE, article=TRUE )
+  {
+ rdf <- subset( rT1$Ad, Ae %in% e.pr )
+ art <- rT1$A0
+ prt <- rT1$Pr
+ crt <- rT1$Cr
+
+ # nice plotting frame for rates
+ plot( NA, log="y",
+       xlim=c(10,85), ylim=rlim, yaxs="i", xlab="", ylab="" )
+ abline( v=e.pr, col=gray(0.8) )
+ axis( side=1, at=seq(10,85,10), labels=NA, tcl=-0.4 )
+ axis( side=1, at=seq(10,85, 5), labels=NA, tcl=-0.2 )
+ axis( side=2, at=outer(2:10,-1:2,function(x,y) x*10^y), labels=NA, tcl=-0.3 )
+ # rates plotted
+ for( e in e.pr )
+    {
+ matshade( rdf$Ax[rdf$Ae==e], rdf[rdf$Ae==e,4:9],
+           type="l", lwd=2, lty=1,
+           col=c("blue","red",gray(0.3)), )
+ text( e+0.5, max(rdf[rdf$Ae==e,4:9][2,],na.rm=TRUE)*tf,
+       paste("Diagn.\nage ",e,sep=""), adj=c(0,0) )
+    }
+ # matshade( art[,1], art[,2:10],
+ matlines( art[,1], art[,c(2,5)],
+           type="l", lwd=3, lty="11", lend="butt",
+           col=c("blue","red",gray(0.5)) )
+ # matshade( art[,1], art[,11:19],
+ matlines( art[,1], art[,c(11,14)],
+           type="l", lwd=1, lty=1, lend="butt",
+           col=c("blue","red",gray(0.5)) )
+  }
> par( mfrow=c(2,1), oma=c(2,2,2,0), mar=c(1,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltpa( rT1, e.pr=1:5*10+5, rlim=c(0.3,250), tf=1 )
> pltpa( rT2, e.pr=3:7*10  , rlim=c(0.3,250), tf=1.1 )
> mtext( "Age at follow-up",           side=1, line=1, outer=TRUE )
> mtext( c("T1D mortality rate per 1000 PY",
+          "T2D mortality rate per 1000 PY"),
+         at=c(3,1)/4, side=2, line=1, outer=TRUE, las=0 )
```

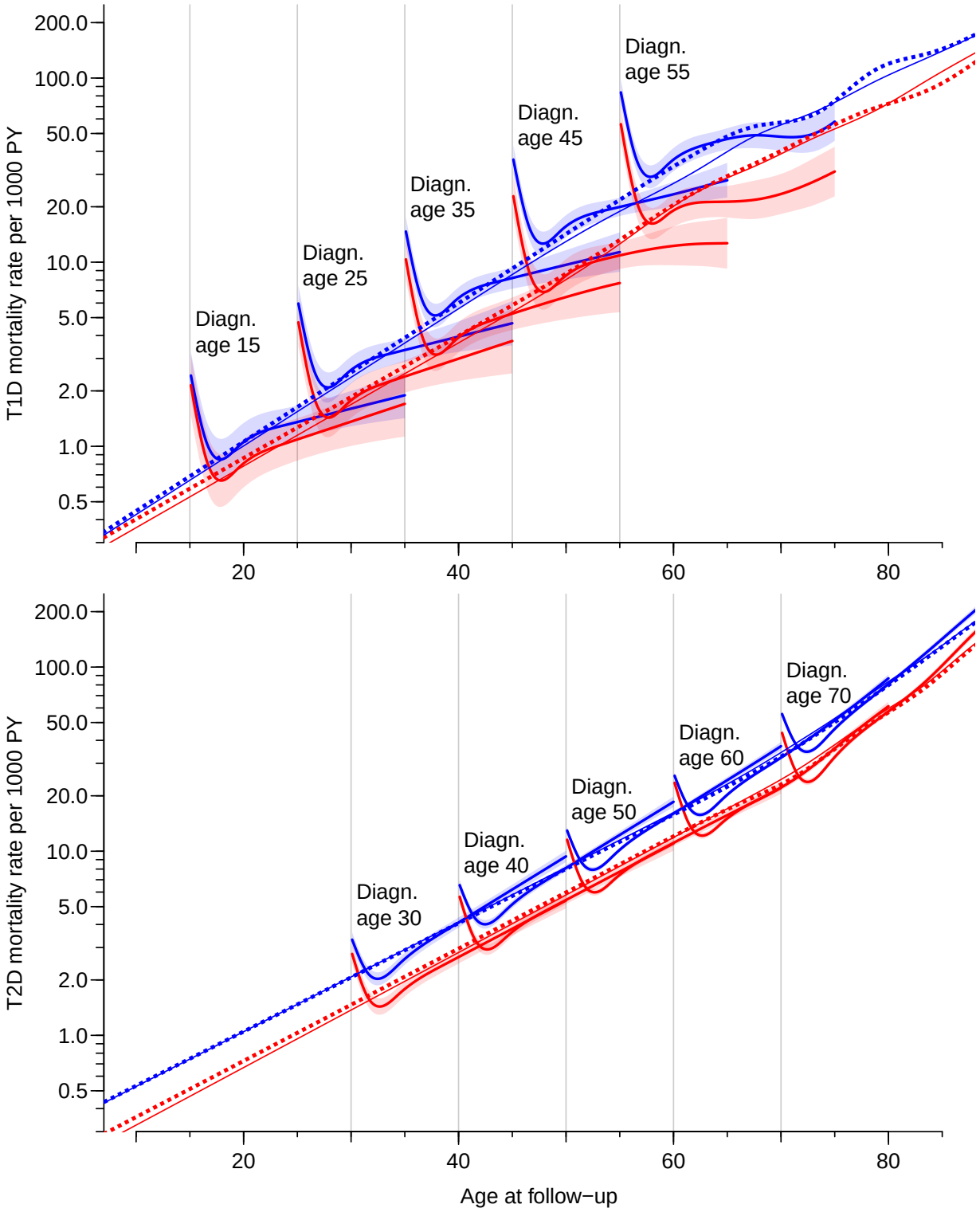


Figure 7.3: *Mortality rates at 2015-01-01 for persons diagnosed in differewnt ages, followed for 0–20 years of diabetes duration. Broken lines are mortality rates modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall mortality also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men; shaded areas indicate 95% confidence intervals.
`./graph/mort-art-m`

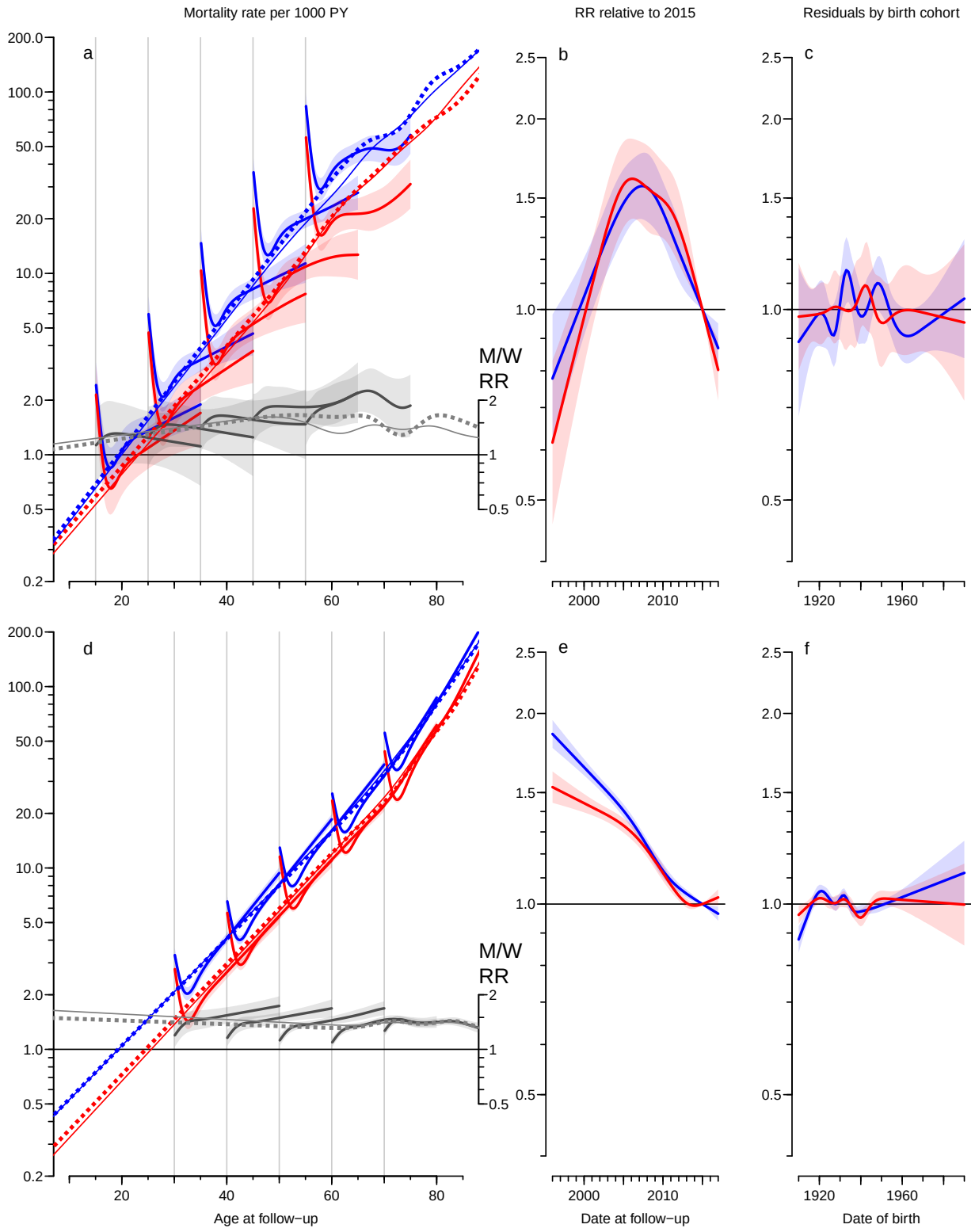


Figure 7.4: *Mortality, and RR relative to 2015-01-01 and birth cohort residuals. Leftmost plot shows the mortality rates at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are mortality rates modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall mortality also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W RR; shaded areas indicate 95% confidence intervals. `./graph/mort-all-m mort.tex`

Finally, we make the same set of plots, but for all diabetes combined (it is going to be quite close to the T2D results)

```
> layout( matrix(1:3,1,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rDM, coh=TRUE, nlt=0, e.pr=15*1:5 )
> mtext( c("Mortality rate per 1000 PY",
+          "RR relative to 2015",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE, cex=0.67 )
> 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 )
```

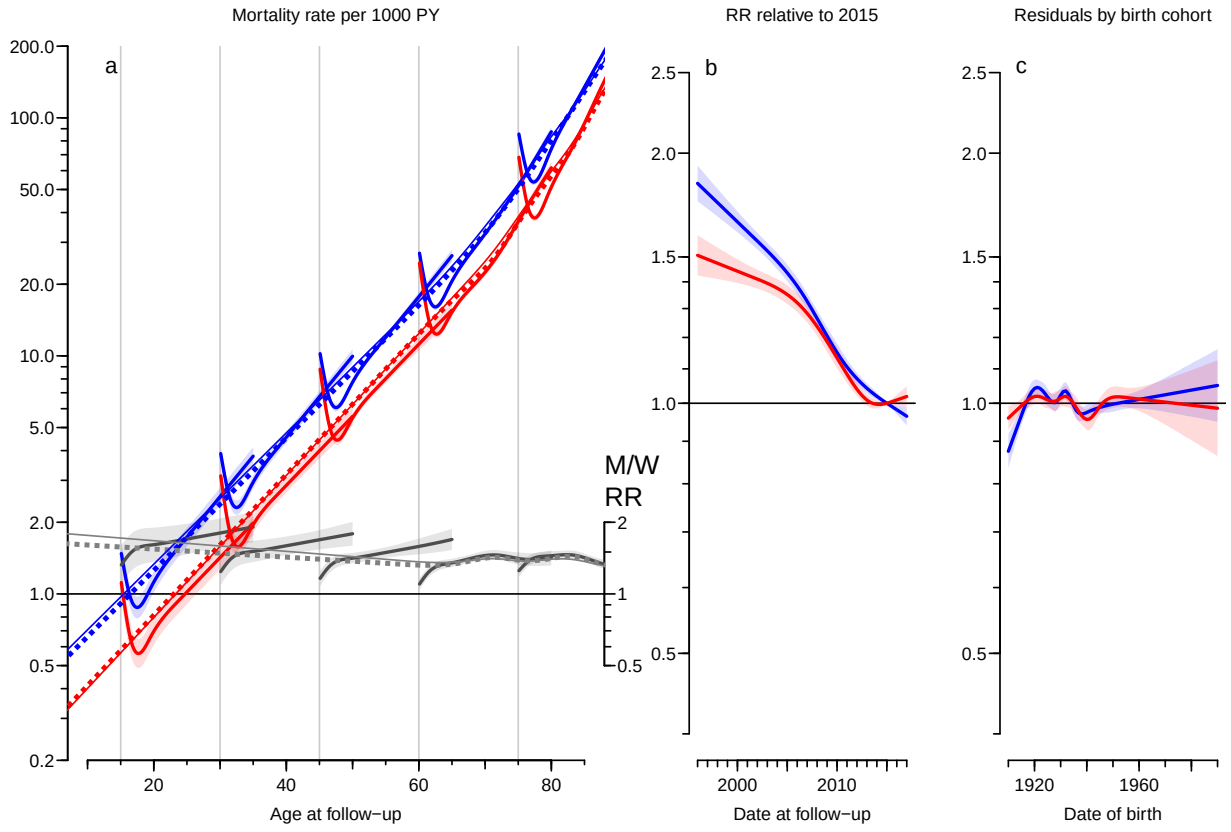


Figure 7.5: *Mortality, HR relative to 2015-01-01 and birth cohort residuals. Left panel shows the mortality rates at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken line in left panel is mortality rates modeled ignoring age at diagnosis and duration of diabetes. Thin full line is overall mortality also including prevalent cases as of 1996-01-01. Right panel is the mortality RR relative to 2015-10-01. Red curves are for women, blue for men, black are M/W RR; shaded areas indicate 95% confidence intervals.* `./graph/mort-DM-m`

For the paper proper we leave out the cohort residuals:

```
> layout( matrix(1:4,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n", cex=0.84 )
> pltp( rDM, coh=FALSE, nlt=0, e.pr=15*1:5, rlim=c(0.1,200), rrpos=0.2 )
> mtext( c("Mortality rate per 1000 PY",
+          "RR relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

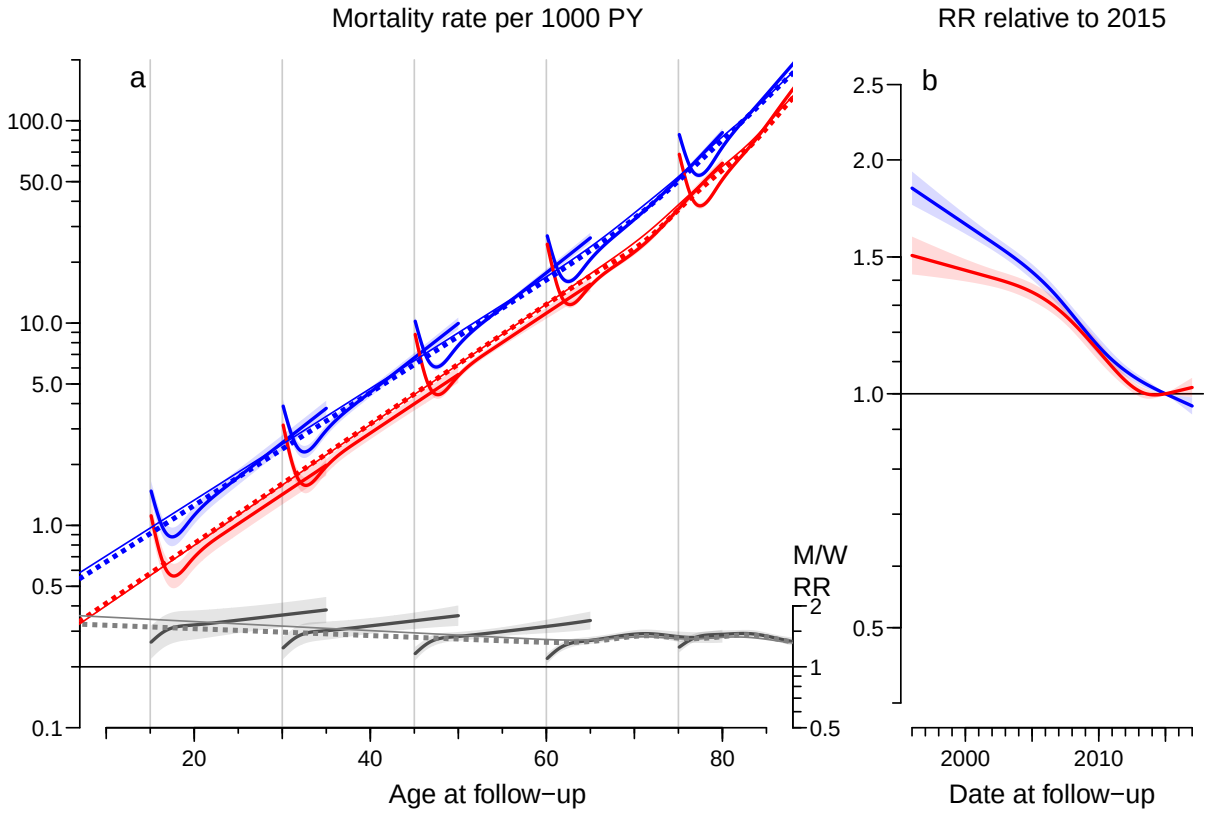


Figure 7.6: *Mortality and mortality RR relative to 2015-01-01. Leftmost plot shows the mortality rates at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are mortality rates modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall mortality also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W RR; shaded areas indicate 95% confidence intervals. `./graph/mort-aDM-m`

7.5 Comparison of T1 and T2 mortality

Since we fitted models separately for T1 and T2 diabetes we can — just as we did for the M/W comparison — derive mortality rate-rations between T1 and T2 mortality.

Specifically, the models we reported from were for current age, a, age at diagnosis e and diabetes duration $d = a - e$ at calendar time p:

$$\log\big(\lambda(a,e,d,p)\big) = f(a) + g(e) + h(d) + j(p)$$

This was fitted separately for each sex and diabetes type, so for a given sex, the T2/T1 mortality rate-ratio is:

$$\log\big(\mathrm{RR}(a,e,d,p)\big) = \big(f_2(a) - f_1(a) + g_2(e) - g_1(e) + h_2(d) - h_1(d)\big) + \big(j_2(p) - j_1(p)\big)$$

So we make a graph similar to the previous ones, but now for the T2/T1 ratios, using each of the two terms in the larger brackets.

We can even produce the M/W ratios of mortality RRs, exploring whether the two types of diabetes have the same impact on mortality for men and women.

Since calculations are done on separate subsets of data, calculation of confidence intervals are straight forward. The 12 variables in `rT1$Ad` are the age/diag/dur terms from the mortality models for T1, the first three variables are age at entry (`Ae`), diabetes duration `dur` and age at follow-up (`Ax`), and the remaining 9 columns are the mortality rates for men, women and the M/W-RR of rates. Thus all we need is to take the ratio of these between T1 and T2 (computing the confidence limits too, of course).

Similar calculations can be done for the structures `A0` ignoring the duration and the `Px` of the secular trends.

```
> RR0 <- rT1$A0
> RRd <- rT1$Ad
> RRp <- rT1$Pr
> for( i in 0:5 ) {wh<-1+i*3+(1:3) ; RR0[,wh] <- ci.ratio( as.matrix(rT1$A0[,wh]),
+                                                          as.matrix(rT2$A0[,wh]) ) }
> for( i in 0:2 ) {wh<-3+i*3+(1:3) ; RRd[,wh] <- ci.ratio( as.matrix(rT1$Ad[,wh]),
+                                                          as.matrix(rT2$Ad[,wh]) ) }
> for( i in 0:1 ) {wh<-1+i*3+(1:3) ; RRp[,wh] <- ci.ratio( as.matrix(rT1$Pr[,wh]),
+                                                          as.matrix(rT2$Pr[,wh]) ) }
```

Finally we put structures in a list so it can be referenced in the function `pltp`:

```
> RR <- list( A0=RR0, Ad=RRd, Pr=RRp )
```

We simply re-use the code for the mortality curves and plot the graphs not of mortality rates, but of T1D vs. T2D RRs:

```
> layout( matrix(1:4,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n", cex=0.83 )
> pltp( RR, coh=FALSE, nlt=0, rlim=c(0.1,5), e.pr=c(30,45,60), rrpos=0.2, rr=TRUE )
> # For illustration a decline of 3\%/year
> x <- 2010:2017 ; y <- 0.97^(x-2015) ; lines(x,y,lty=3)
> mtext( c("T1D vs. T2D mortality rate-ratio at 2015",
+          "RR ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date of follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

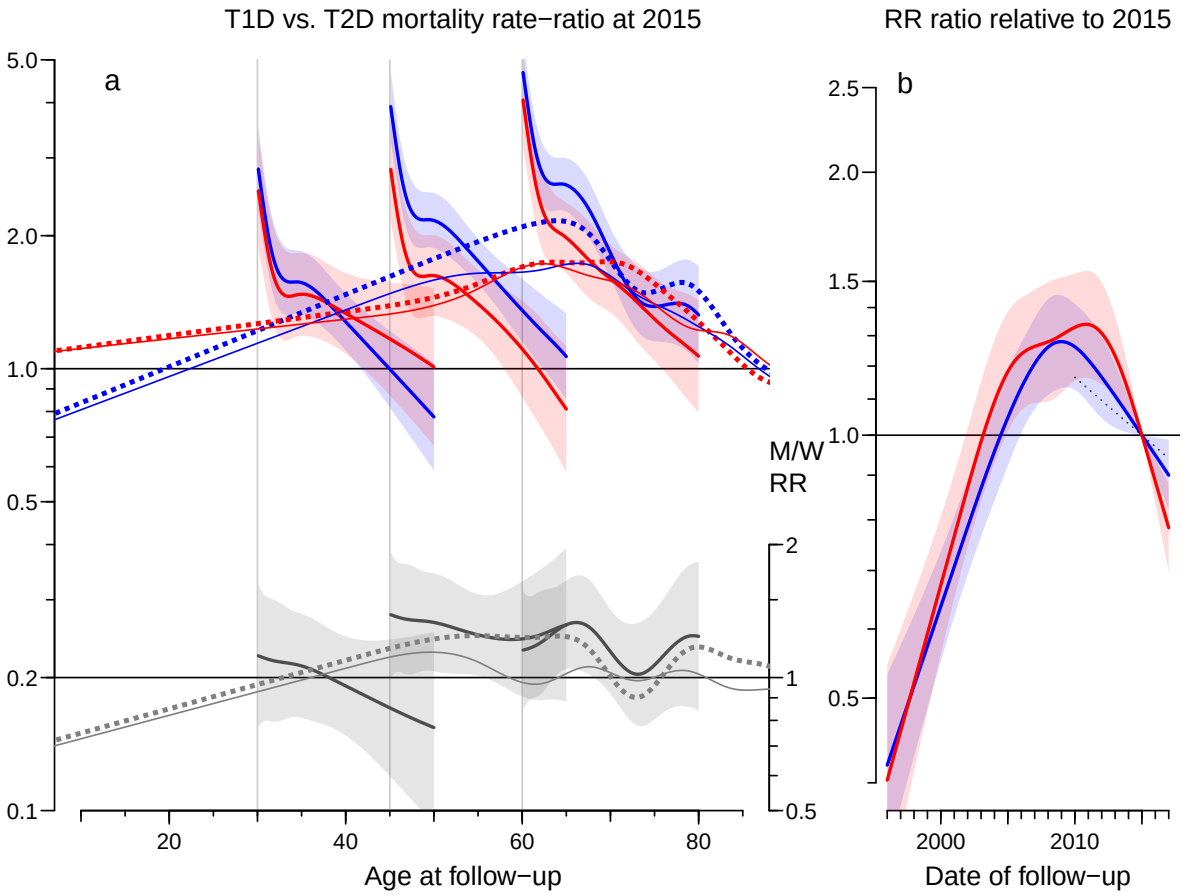


Figure 7.7: *T1D versus T2D mortality RR at 2015-01-01. Leftmost plot shows the mortality RR at 2015-01-01 for persons diagnosed in different ages.*
Broken lines in leftmost plot are mortality RRs modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall mortality RR also including prevalent cases as of 1996-01-01.
Red curves are for women, blue for men, black are M/W RR ratio; shaded areas indicate 95% confidence intervals. `./graph/mort-t1t2-rr`

From figure 8.7 we see that in the range of age at diagnosis 40–60, the T1D vs T2D mortality rate ratio is broadly 2 for women and 2.5 for men during the first few years after diagnosis, and after that attenuates by duration / age. Ignoring duration of diabetes and age at diagnosis, showed a T1/T2 rate ratio of 1.5 for man and 1.2 for women at age 50. In the range 30–70 years of age the T1/T2 mortality RR was higher among men than among women; outside this interval there was essentially no detectable difference in RRs between man and women.

The trend in RR in the right panel shows that the RR among T1D patients have been declining by some 3% per year since 2010 (direct measurement on the graph, the broken black line has a slope of -3%/year).

7.5.1 Direct comparison of T1D vs. T2D mortality

We can derive anm overall mortality RR between T1D and T2D by fitting a joint model for T1D and T2D pateinst taking current age, duration, at at diagnosis and period into account:

```
> rtD <-  subset( rtL, state!="noDM" )
> a.kn <-     with( rtD, quantile( rep(Ax    ,D), (1:8-0.5)/8 ) )
> d.kn <- c(0,with( rtD, quantile( rep(   dur,D), (1:3-0.0)/4 ) ))
> e.kn <-     with( rtD, quantile( rep(Ax-dur,D), (1:5-0.5)/5 ) )
> p.kn <-     with( rtD, quantile( rep(   Px ,D), (1:5-0.5)/5 ) )
> c.kn <-     with( rtD, quantile( rep(Px-Ax ,D), (1:9-0.5)/9 ) )
> b12 <- glm( D ~ Ns( Ax    , knots=a.kn ) +
+                 Ns(    dur, knots=d.kn ) +
+                 Ns( Ax-dur, knots=e.kn ) +
+                 Ns( Px    , knots=p.kn ) +
+                 Ns( Cx    , knots=c.kn ) + state + sex,
+             offset = log(Y),
+             family = poisson,
+               data = rtD )
> m12 <- update( b12, . ~ . - sex, data = subset( rtD, sex=="M" ) )
> w12 <- update( b12, . ~ . - sex, data = subset( rtD, sex=="W" ) )
> round( cbind(
+        ci.exp( m12, subset="st", ctr.mat=rbind( 1, -1) ),
+        ci.exp( w12, subset="st", ctr.mat=rbind( 1, -1) ),
+        ci.exp( b12, subset="st", ctr.mat=rbind( 1, -1) ) ), 2 )
     exp(Est.) 2.5% 97.5% exp(Est.) 2.5% 97.5% exp(Est.) 2.5% 97.5%
[1,]      0.54 0.52  0.56      0.64 0.61  0.68      0.58 0.56  0.60
[2,]      1.86 1.79  1.93      1.55 1.48  1.63      1.73 1.68  1.78
```

Thus we see that there is an overall different T1 / T2 mortality rate ratio between man and women, for man it is 1.86, for women 1.55. This is controlled for age at FU, duration, age at diagnosis, calendar time at FU, and date of birth.

Trend in T1 / T2 mortality rate ratio

To produce an overall assessment of trens we fit a simple model where we model the time trend in the mortality rate ratio:

```
> b12p <- update( b12, . ~ . - Ns( Px, knots=p.kn )
+                            - Ns( Cx, knots=c.kn )
+                            + I(Px-2010):state )
> m12p <- update( b12p, . ~ . - sex, data = subset( rtD, sex=="M" ) )
> w12p <- update( b12p, . ~ . - sex, data = subset( rtD, sex=="W" ) )
> round( ci.exp( b12p ), 3 )
```

```
                             exp(Est.)   2.5%  97.5%
(Intercept)                     48.689 46.793 50.661
Ns(Ax, knots = a.kn)1            2.402  2.293  2.517
Ns(Ax, knots = a.kn)2            3.626  3.464  3.796
Ns(Ax, knots = a.kn)3            5.164  4.922  5.417
Ns(Ax, knots = a.kn)4            7.303  6.987  7.633
Ns(Ax, knots = a.kn)5           10.234  9.879 10.602
Ns(Ax, knots = a.kn)6           25.707 24.307 27.187
Ns(Ax, knots = a.kn)7           14.776 14.395 15.167
Ns(dur, knots = d.kn)1           0.690  0.677  0.704
Ns(dur, knots = d.kn)2           0.245  0.234  0.257
Ns(dur, knots = d.kn)3           0.895  0.880  0.911
Ns(Ax - dur, knots = e.kn)1      1.146  1.109  1.184
Ns(Ax - dur, knots = e.kn)2      1.093  1.067  1.120
Ns(Ax - dur, knots = e.kn)3      1.154  1.102  1.209
Ns(Ax - dur, knots = e.kn)4      1.000  1.000  1.000
stateT2                          0.571  0.553  0.589
sexW                             0.708  0.699  0.717
stateT1:I(Px - 2010)             0.977  0.971  0.982
stateT2:I(Px - 2010)             0.971  0.970  0.972
```

```
> rch <- cbind(
+ ci.exp( m12p, subset="Px", ctr.mat=rbind(1:0,0:1,c(1,-1)) ),
+ ci.exp( w12p, subset="Px", ctr.mat=rbind(1:0,0:1,c(1,-1)) ),
+ ci.exp( b12p, subset="Px", ctr.mat=rbind(1:0,0:1,c(1,-1)) ) )
> colnames( rch )[c(1,4,7)] <- c("M","W","M+W")
> rownames( rch ) <- c("T1D","T2D","T1D/T2D")
> round( rch, 3 )
```

```
            M  2.5% 97.5%     W  2.5% 97.5%   M+W  2.5% 97.5%
T1D     0.974 0.967 0.980 0.981 0.973 0.990 0.977 0.971 0.982
T2D     0.968 0.966 0.969 0.975 0.973 0.977 0.971 0.970 0.972
T1D/T2D 1.006 0.999 1.013 1.006 0.997 1.015 1.006 1.000 1.011
```

```
> round( (rch-1)*100, 1 )
```

```
           M 2.5% 97.5%    W 2.5% 97.5%  M+W 2.5% 97.5%
T1D     -2.6 -3.3  -2.0 -1.9 -2.7  -1.0 -2.3 -2.9  -1.8
T2D     -3.2 -3.4  -3.1 -2.5 -2.7  -2.3 -2.9 -3.0  -2.8
T1D/T2D  0.6 -0.1   1.3  0.6 -0.3   1.5  0.6  0.0   1.1
```

7.6 SMR

The analysis of SMR is parallel to the analysis of rates, except that we replace the person-years by the expected number of events. This is however easily done by creating a new dataset with expected numbers; first by extraction the rates (per 1000 PY) from the noDM part of data:

```
> names(rt)
 [1] "sex"   "state" "A"     "P"     "C"     "dur"   "Dcvd"  "Dcan"  "Dres"  "Doth"  "Y"
[12] "T1"    "T2"    "D"     "Ax"    "Px"    "Cx"
```

```
> nDM <- transform( subset( rt,
+                           state=="noDM",
+                           select=c("sex","Ax","Px","D","Y") ),
+                   R = D / Y )[,c("sex","Ax","Px","R")]
> table( with( nDM, table(sex,Ax,Px) ) )
```

```
   0    1
8400 8400
```

```
> table( with( nDM, table(sex,floor(Ax),Px) ) )
```

```
   1
8400
```

```
> table( with( nDM, table(sex,floor(Ax),floor(Px)) ) )
```

```
   2
4200
```

— then merge with the original data computing the expected numbers. However, first note that it is immaterial what units `Y` is measured in, as long as it is in the same units across the dataset `rt`. Second, note that we are naming the variable with expected numbers "`Y`", so we can use the same code as before:

```
> xt  <- transform( merge( rt , nDM, all=TRUE ),
+                   Y = Y * R )
> xtL <- transform( merge( rtL, nDM, all=TRUE ),
+                   Y = Y * R )
> # check that Obs=Exp in the no DM group
> round( do.call( rbind, with( xt, tapply( D-Y, state, range ) ) ), 4 )
        [,1]    [,2]
noDM  0.0000  0.0000
T1   -1.0414  5.4500
T2   -4.4204 20.0156
```

Now use the `mort.res` function to generate SMR results for T1, T2 and combined DM:

```
> xT1 <- mort.res( subset( xtL, state=="T1" & Y>0 ),
+                  subset( xt , state=="T1" & Y>0 ) )
> xT2 <- mort.res( subset( xtL, state=="T2" & Y>0 ),
+                  subset( xt , state=="T2" & Y>0 ) )
> xDM <- mort.res( subset( xtL, state %in% c("T1","T2") & Y>0 ),
+                  subset( xt , state %in% c("T1","T2") & Y>0 ) )
> save( xT1, xT2, xDM, file="../data/smr-res.Rda" )
```

An overview of the annual changes in SMR:

```
> load( file="../data/smr-res.Rda" )
> fCp( xT1$tr, d=3, z=NULL )
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.000     0.000     2.549     3.416     1.690
W     0.000     0.000     2.450     3.511     1.400
```

```
> fCp( xT2$tr, d=3, z=NULL )
```

```
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.000     0.000    -0.461    -0.280    -0.641
W     0.000     0.027    -0.231    -0.027    -0.435
```

```
> fCp( xDM$tr, d=3, z=NULL )
```

```
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.000     0.000    -0.542    -0.366    -0.718
W     0.000     0.010    -0.263    -0.063    -0.463
```

7.6.1 Plotting the SMR

With the function `pltp` we plot a 2 by 3 layout aimed at the ESM:

```
> layout( matrix(1:6,2,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,1), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xT1, coh=TRUE, nlt=0, rlim=c(0.3,20), e.pr=1:5*10+5 )
> pltp( xT2, coh=TRUE, nlt=3, rlim=c(0.3,20), e.pr=3:7*10 )
> mtext( c("SMR",
+          "SMR ratio relative to 2015",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE, cex=0.67 )
> 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 )
```

For the paper proper we leave out the cohort residuals:

```
> par( mfrow=c(2,1), oma=c(2,2,2,0), mar=c(1,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltpa( xT1, rlim=c(0.9,25), e.pr=1:5*10+5, tf=0.5 )
> abline( h=1, col=gray(0.8) )
> pltpa( xT2, rlim=c(0.9,25), e.pr=3:7*10, tf=1   )
> abline( h=1, col=gray(0.8) )
> mtext( "Age at follow-up", side=1, line=1, outer=TRUE )
> mtext( c("T1D SMR", "T2D SMR"),
+          at=c(3,1)/4, side=2, line=1, outer=TRUE, las=0 )
```

. . . and separate versions for slides:

```
> layout( matrix(1:2,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(1,1,2,0), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xT1, coh=FALSE, lab="T1D", nlt=0, rlim=c(0.3,20), e.pr=1:5*10+5 )
> mtext( c("SMR at 2015",
+          "SMR ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )

> layout( matrix(1:2,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(1,1,2,0), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xT2, coh=FALSE, lab="T2D", nlt=2, rlim=c(0.3,20), e.pr=3:7*10 )
> mtext( c("SMR at 2015",
+          "SMR ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

Finally we make the same set of plots, but for all diabetes combined (it is going to be quite close to the T2D results)

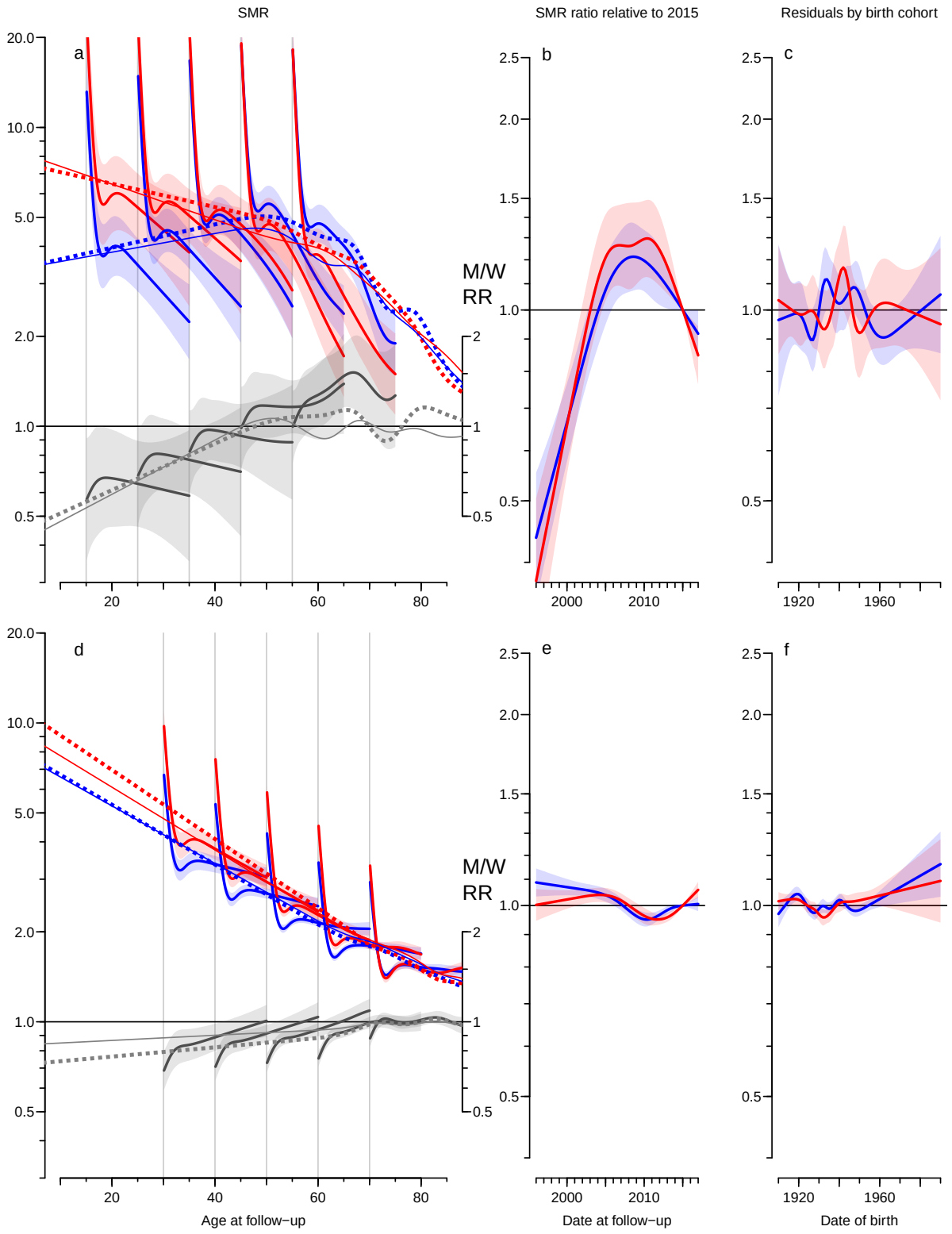


Figure 7.8: *SMR and SMR-ratio relative to 2015-01-01 and birth cohort residuals. Leftmost plot shows the SMR at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are SMR modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall SMR also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W SMR ratios; shaded areas indicate 95% confidence intervals. `./graph/mort-all-smr`
`mort.tex`

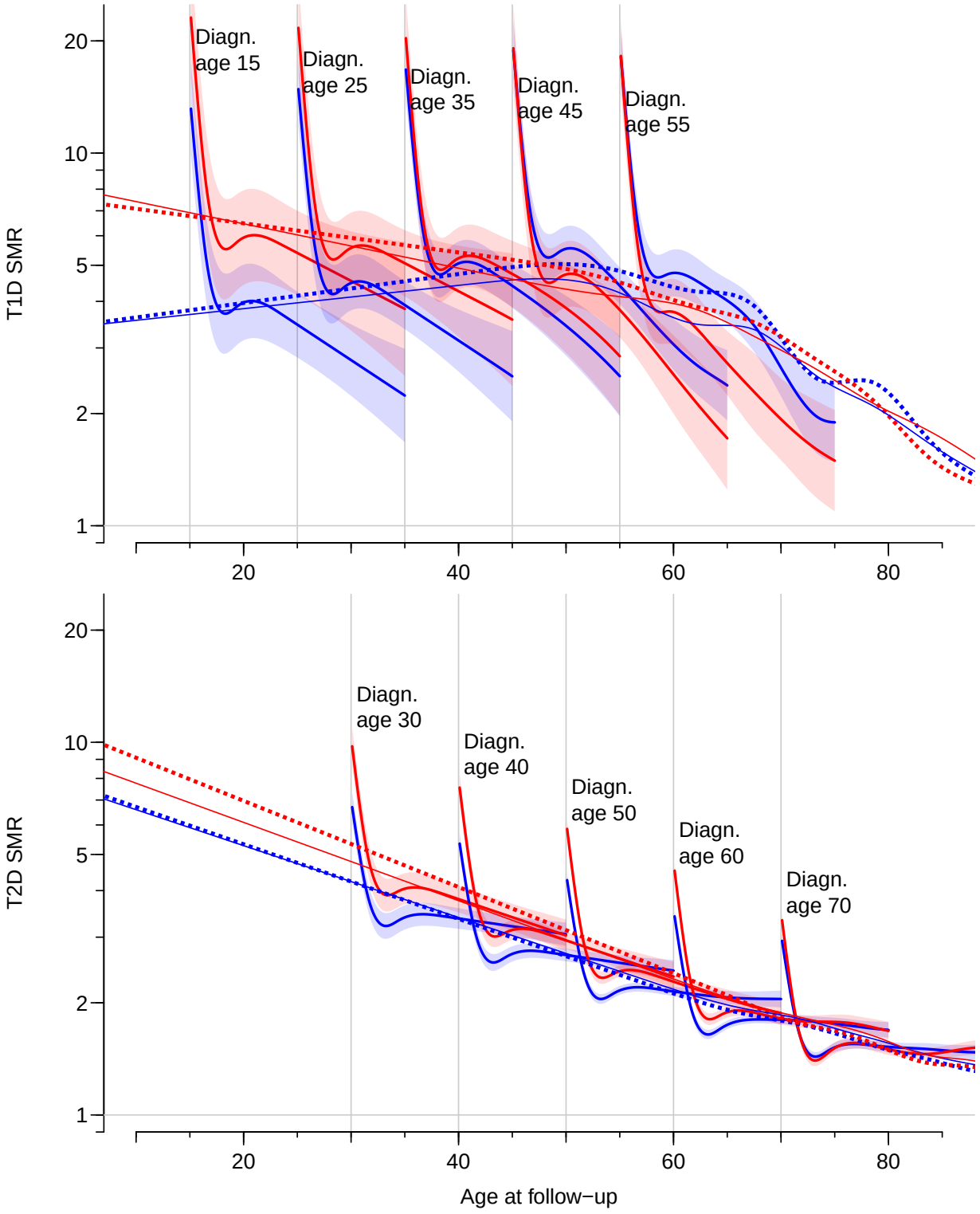


Figure 7.9: *SMR and SMR-ratio relative to 2015-01-01. Leftmost plot shows the SMR at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are SMR modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall SMR also including prevalent cases as of 1996-01-01. Red curves are for women, blue for men, black are M/W SMR ratios; shaded areas indicate 95% confidence intervals.* `./graph/mort-art-smr`

```
> layout( matrix(1:3,1,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xDM, coh=TRUE, lab="DM", nlt=0, rlim=c(0.3,30), e.pr=c(15,30,45,60,75) )
> mtext( c("SMR",
+          "SMR ratio relative to 2015",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE, cex=0.84 )
> mtext( c("Age at follow-up",
+          "Date at follow-up",
+          "Date of birth"),
+        at = c(4,10,14)/16,
+        side=1, outer=TRUE )
```

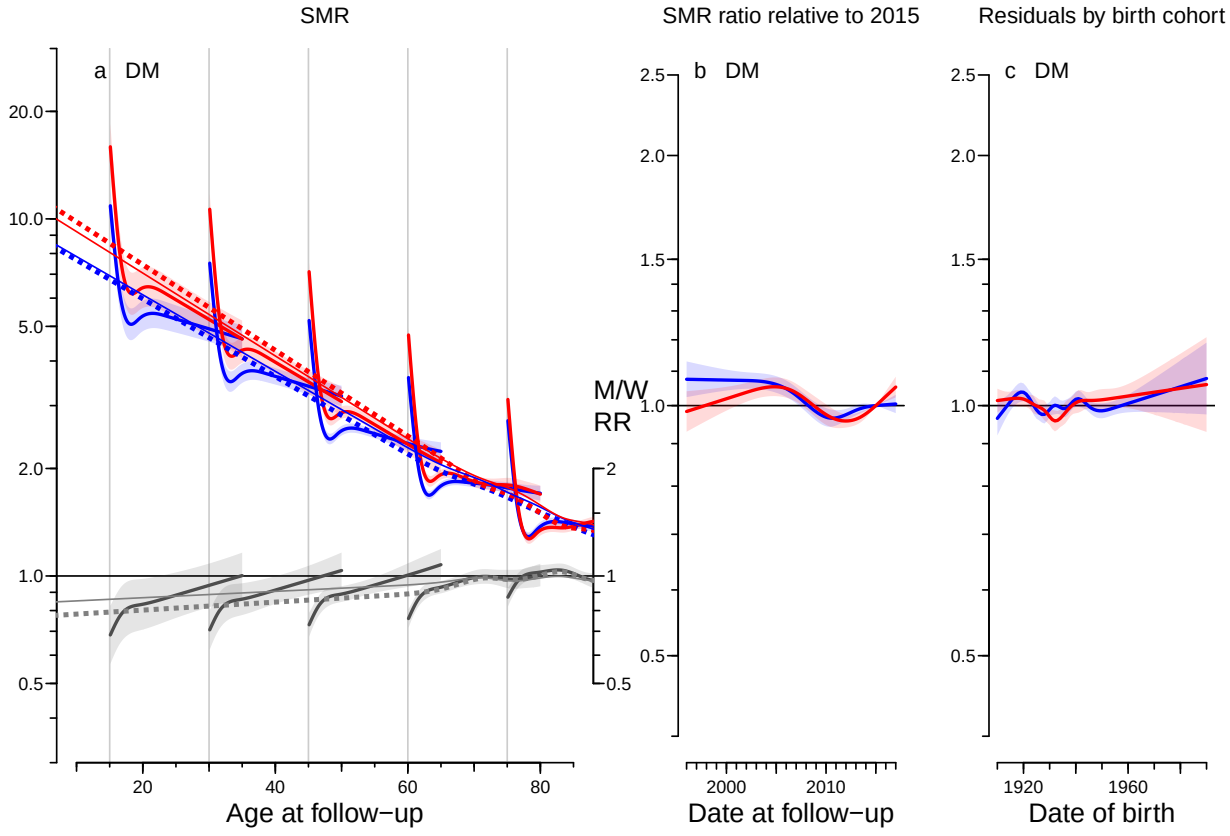


Figure 7.10: *SMR and SMR-ratio for all DM, relative to 2015-01-01 and birth cohort residuals. Leftmost plot shows the SMR at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are SMR modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall SMR also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W SMR ratios; shaded areas indicate 95% confidence intervals. `./graph/mort-DM-smr`

For the paper proper we leave out the cohort residuals:

```
> layout( matrix(1:4,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
```

```
> pltp( xDM, coh=FALSE, lab="DM", nlt=0, rlim=c(0.3,20), e.pr=c(15,30,45,60,75) )
> mtext( c("SMR",
+          "SMR ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

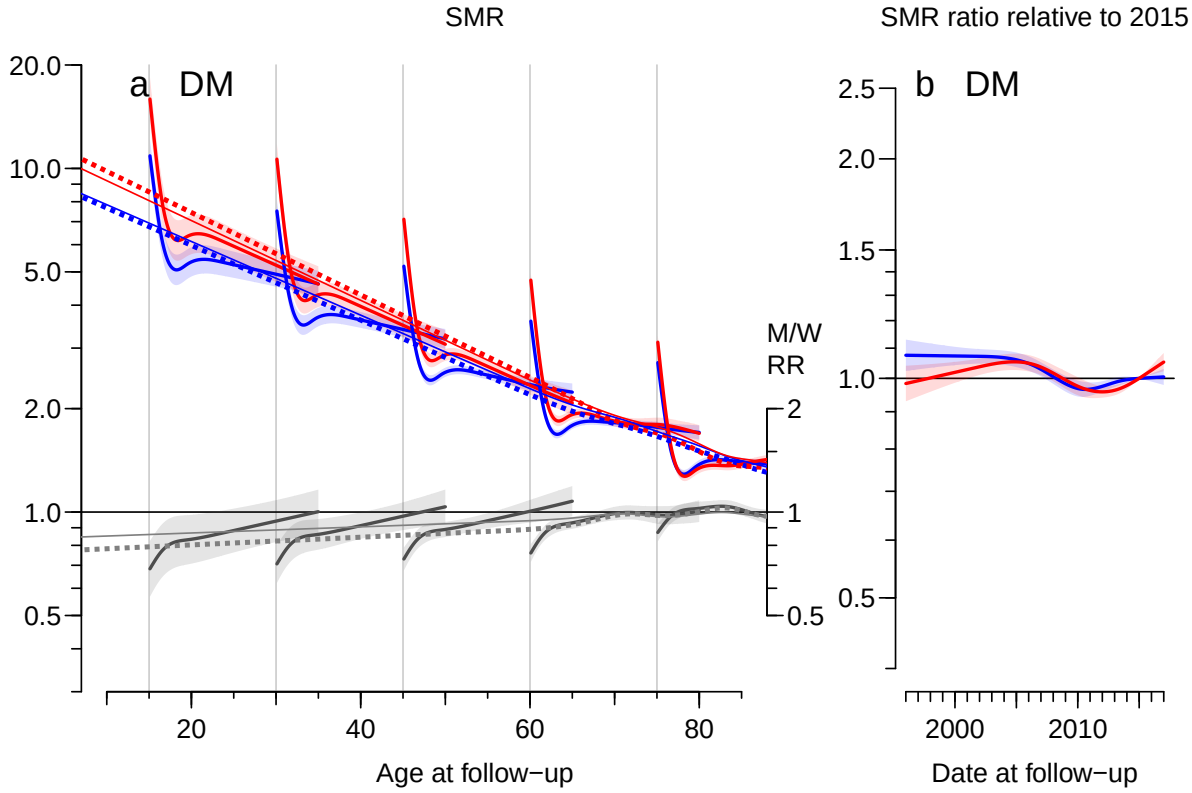


Figure 7.11: *SMR and SMR-ratio for all DM, relative to 2015-01-01. Leftmost plot shows the SMR at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are SMR modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall SMR also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W SMR ratios; shaded areas indicate 95% confidence intervals. `./graph/mort-aDM-smr`

```
----------------------
2019-01-10 at 17:57:17
Time elapsed: 00:01:35
----------------------
```

. . . now input from `mortx.tex`

Chapter 8

Mortality after 2005

The classification of T1D, and most pronounced presumably of deaths among T1D patients, is less reliable before 2005. By that token we repeat the entire analusis of mortality restricted to follow-up after 2005.

First we load packages and the rate datasets:

```
> library( Epi )
> library( splines )
> library( mgcv )
> start()
----------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-01-10 18:10:02
----------------------------------------
> load( file="../nydata/inits.Rda" )
> load( file="../nydata/rt.Rda" )
```

Havin loaded the rate data, we restrict to the period afte 2005-01-01, and run the same code as before (except that we later also restrict the calendar time period in the prediction frames):

```
> rt  <- subset( rt , Px>2005 )
> rtL <- subset( rtL, Px>2005 )
```

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.

First we summarize the ages at death — the column `mod-80` gives the width of the age-range from the 10^{th} to the 90^{th} percentile of ages at death:

```
> dthA <- do.call( rbind,
+                  with( rt,
+                        tapply( rep(Ax,D),
+                                rep(state,D),
+                                quantile,
+                                c(1,5,9)/10 ) ) )
> dthA <- cbind( dthA, dthA[,3]-dthA[,1] )
> colnames( dthA )[4] <- "mid80"
> round( dthA, 1 )
      10%  50%  90% mid80
noDM 57.3 79.7 92.7  35.3
T1   48.3 70.3 86.7  38.3
T2   62.7 78.7 90.7  28.0
```

Not surprisingly the T2 deaths occur in a much narrower intervals, mainly because the persons at risk are much older.

8.1 Mortality overview

For convenience we make special data frames of T1 follow-up, one with and one without duration, but first an overview of deaths ex- resp. in-cluding the prevalent cases as of 1996-01-01:

```
> # Only persons diagnosed after 1996-01-01
> mL <- xtabs( D ~ floor(P) + state + sex, data=rtL )
> fCtable( addmargins(mL[,c(2,3,1),],2:3),w=6, col.vars=2:3 )
```

state	T1			T2			noDM			Sum		
sex	M	W	Sum	M	W	Sum	M	W	Sum	M	W	
floor(P)												
2005	155	86	241	2,509	2,081	4,590	22,571	24,258	46,829	25,235	26,425	51
2006	149	94	243	2,645	2,194	4,839	22,629	23,987	46,616	25,423	26,275	51
2007	141	90	231	2,895	2,382	5,277	22,587	24,433	47,020	25,623	26,905	52
2008	165	89	254	2,965	2,410	5,375	22,169	23,497	45,666	25,299	25,996	51
2009	156	95	251	3,398	2,647	6,045	21,947	23,638	45,585	25,501	26,380	51
2010	157	87	244	3,540	2,885	6,425	21,709	22,949	44,658	25,406	25,921	51
2011	127	79	206	3,749	2,875	6,624	20,768	22,092	42,860	24,644	25,046	49
2012	124	70	194	4,132	3,022	7,154	20,306	21,793	42,099	24,562	24,885	49
2013	125	52	177	4,289	3,238	7,527	20,336	21,411	41,747	24,750	24,701	49
2014	97	68	165	4,640	3,477	8,117	19,773	20,703	40,476	24,510	24,248	48
2015	95	49	144	4,783	3,721	8,504	20,345	21,169	41,514	25,223	24,939	50
2016	107	53	160	5,006	3,789	8,795	20,112	21,078	41,190	25,225	24,920	50

```
> # All persons
> mA <- xtabs( D ~ floor(P) + state + sex, data=rt  )
> fCtable( addmargins(mA[,-1,],1:3),w=6, col.vars=2:3 )
```

state	T1			T2			Sum		
sex	M	W	Sum	M	W	Sum	M	W	Sum
floor(P)									
2005	426	285	711	3,538	3,190	6,728	3,964	3,475	7,439
2006	397	291	688	3,637	3,201	6,838	4,034	3,492	7,526
2007	414	281	695	3,732	3,334	7,066	4,146	3,615	7,761
2008	381	258	639	3,788	3,220	7,008	4,169	3,478	7,647
2009	381	259	640	4,254	3,434	7,688	4,635	3,693	8,328
2010	375	235	610	4,241	3,629	7,870	4,616	3,864	8,480
2011	339	211	550	4,432	3,550	7,982	4,771	3,761	8,532
2012	286	173	459	4,791	3,666	8,457	5,077	3,839	8,916
2013	281	142	423	4,902	3,829	8,731	5,183	3,971	9,154
2014	222	146	368	5,198	3,966	9,164	5,420	4,112	9,532
2015	204	129	333	5,303	4,243	9,546	5,507	4,372	9,879
2016	202	140	342	5,484	4,257	9,741	5,686	4,397	10,083
Sum	3,908	2,550	6,458	53,300	43,519	96,819	57,208	46,069	103,277

```
> fCtable( addmargins(mA[,1,,drop=F],c(1,3)),w=6, col.vars=2:3 )
```

state	noDM		
sex	M	W	Sum
floor(P)			
2005	22,571	24,258	46,829
2006	22,629	23,987	46,616
2007	22,587	24,433	47,020
2008	22,169	23,497	45,666

```
2009           21,947  23,638  45,585
2010           21,709  22,949  44,658
2011           20,768  22,092  42,860
2012           20,306  21,793  42,099
2013           20,336  21,411  41,747
2014           19,773  20,703  40,476
2015           20,345  21,169  41,514
2016           20,112  21,078  41,190
Sum           255,252 271,008 526,260
```

The analysis datasets for T1D mortality:

```
> rt1 <- subset( rtL, state=="T1" )
> rx1 <- subset( rt , state=="T1" )
> str( rt1 )
'data.frame':        64058 obs. of  17 variables:
 $ sex  : Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ state: Factor w/ 3 levels "noDM","T1","T2": 2 2 2 2 2 2 2 2 2 2 ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ P    : num  2005 2005 2006 2006 2006 ...
 $ C    : num  2004 2004 2005 2005 2006 ...
 $ dur  : num  0.1 0.35 0.1 0.35 0.1 0.1 0.35 0.75 0.1 0.1 ...
 $ Dcvd : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dcan : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dres : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Doth : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y    : num  9.05e-05 1.81e-04 2.50e-04 2.00e-04 9.79e-05 ...
 $ T1   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Ax   : num  0.667 0.667 0.667 0.667 0.333 ...
 $ Px   : num  2005 2005 2006 2006 2007 ...
 $ Cx   : num  2005 2005 2006 2006 2006 ...

> str( rx1 )

'data.frame':        108873 obs. of  17 variables:
 $ sex  : Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ state: Factor w/ 3 levels "noDM","T1","T2": 2 2 2 2 2 2 2 2 2 2 ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ P    : num  2005 2005 2006 2006 2006 ...
 $ C    : num  2004 2004 2005 2005 2006 ...
 $ dur  : num  0.1 0.35 0.1 0.35 0.1 0.1 0.35 0.75 0.1 0.1 ...
 $ Dcvd : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dcan : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dres : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Doth : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y    : num  9.05e-05 1.81e-04 2.50e-04 2.00e-04 9.79e-05 ...
 $ T1   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Ax   : num  0.667 0.667 0.667 0.667 0.333 ...
 $ Px   : num  2005 2005 2006 2006 2007 ...
 $ Cx   : num  2005 2005 2006 2006 2006 ...
```

8.2 Type 1 diabetes patient mortality

In these 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 for men and women:

```
> ( 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%
37.66667 52.33333 60.33333 65.33333 70.66667 76.66667 81.66667 88.66667
> ( d.kn <- c(0,with( rt1, quantile( rep(   dur,D), (1:3-0.0)/4 ) )) )
      25%  50%  75%
 0.0  3.5  7.5 10.5
> ( e.kn <-     with( rt1, quantile( rep(Ax-dur,D), (1:5-0.5)/5 ) )  )
     10%      30%      50%      70%      90%
35.30167 52.16667 60.83333 69.16667 79.16667
> ( p.kn <-     with( rt1, quantile( rep(   Px ,D), (1:5-0.5)/5 ) )  )
     10%      30%      50%      70%      90%
2006.333 2008.333 2010.333 2012.333 2015.333
> ( 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%
 1920.333  1926.667  1932.333  1938.333  1942.667  1947.333  1952.333  1960.333  1975.074
```

With these knot-vectors in place we set up separate mortality models for men and women, simple age-period models, as well as more elaborate models taking age at diagnosis and duration of diabetes into account.

```
> 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=="W" ) )
> mtx <- update( mt0, data = subset( rx1, sex=="M" ) )
> ftx <- update( mt0, data = subset( rx1, sex=="W" ) )
> 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=="W" ) )
> rt1$Ae <- rt1$Ax - rt1$dur
> ml1 <- update( mt1, . ~ . - Ns( Px, knots=p.kn ) + Px )
> fl1 <- update( ft1, . ~ . - Ns( Px, knots=p.kn ) + Px )
> round( ( rbind( "M"=ci.exp( ml1, subset="Px" )[1,],
+                 "W"=ci.exp( fl1, subset="Px" )[1,] ) - 1 )*100, 1 )
  exp(Est.) 2.5% 97.5%
M      -5.1 -6.6  -3.5
W      -5.0 -7.0  -2.9
```

So we see there is an overall *in*crease in mortality among T1D patients.

As a final check we fit a cohort-effect to the residuals from the age, date duration and period model, to check if there is a residual cohort-effect:

```
> prAP <- predict(mt1,type="response")
> mr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(prAP),
+             family = poisson,
+               data = subset(rt1,sex=="M") )
> prAP <- predict(ft1,type="response")
> fr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(prAP),
+             family = poisson,
+               data = subset(rt1,sex=="W") )
```

We then set up data frames for predicting the mortality rates for men and women as of 2015-01-01, for different combinations of age at diagnosis (entry `e.pr`), date of diagnosis 2000 and a sequence of durations:

```
> e.pr <- seq(15,65,5)
> d.pr <- c(NA,seq(0,20,0.1))
> nd0 <- ndx <- data.frame( Ax = seq( 5,90,0.2), Px=2015, Y=1 )
> nd  <- data.frame( expand.grid( dur=d.pr, Ae=e.pr ) )
> nd <- transform( nd, Ax = Ae+dur,
+                      Px = 2015,
+                       Y = 1 )
> head( nd )
  dur Ae   Ax   Px Y
1  NA 15   NA 2015 1
2 0.0 15 15.0 2015 1
3 0.1 15 15.1 2015 1
4 0.2 15 15.2 2015 1
5 0.3 15 15.3 2015 1
6 0.4 15 15.4 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` (not for matshade till `Epi_2.32`, though):

```
> t0pr <- cbind( ci.pred( mt0, nd0 ), ci.pred( ft0, nd0 ) )
> t0pr <- cbind( t0pr, ci.ratio(t0pr[,1:3],t0pr[,1:3+3]) )
> txpr <- cbind( ci.pred( mtx, ndx ), ci.pred( ftx, ndx ) )
> txpr <- cbind( txpr, ci.ratio(txpr[,1:3],txpr[,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)",
+       xlim=c(10,85), ylim=c(0.2,200), yaxs="i" )
> abline( v=e.pr, col=gray(0.8) )
> abline( h=1 )
> axis( side=1, at=seq(10,85,10), labels=NA, tcl=-0.4 )
> axis( side=1, at=seq(10,85, 5), labels=NA, tcl=-0.2 )
> axis( side=2, at=outer(2:10,-1:2,function(x,y) x*10^y), labels=NA, tcl=-0.3 )
> for( e in e.pr[] )
+ matshade( nd$Ax[nd$Ae==e], t1pr[nd$Ae==e,],
+           type="l", lwd=2, lty=1,
+           col=c("blue","red",gray(0.3)), )
> matshade( nd0$Ax, t0pr,
+           type="l", lwd=3, lty="11", lend="butt",
```

```
+           col=c("blue","red",gray(0.5)) )
> matshade( ndx$Ax, txpr,
+           type="l", lwd=1, lty=1,
+           col=c("blue","red",gray(0.5)) )
> 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 )
```

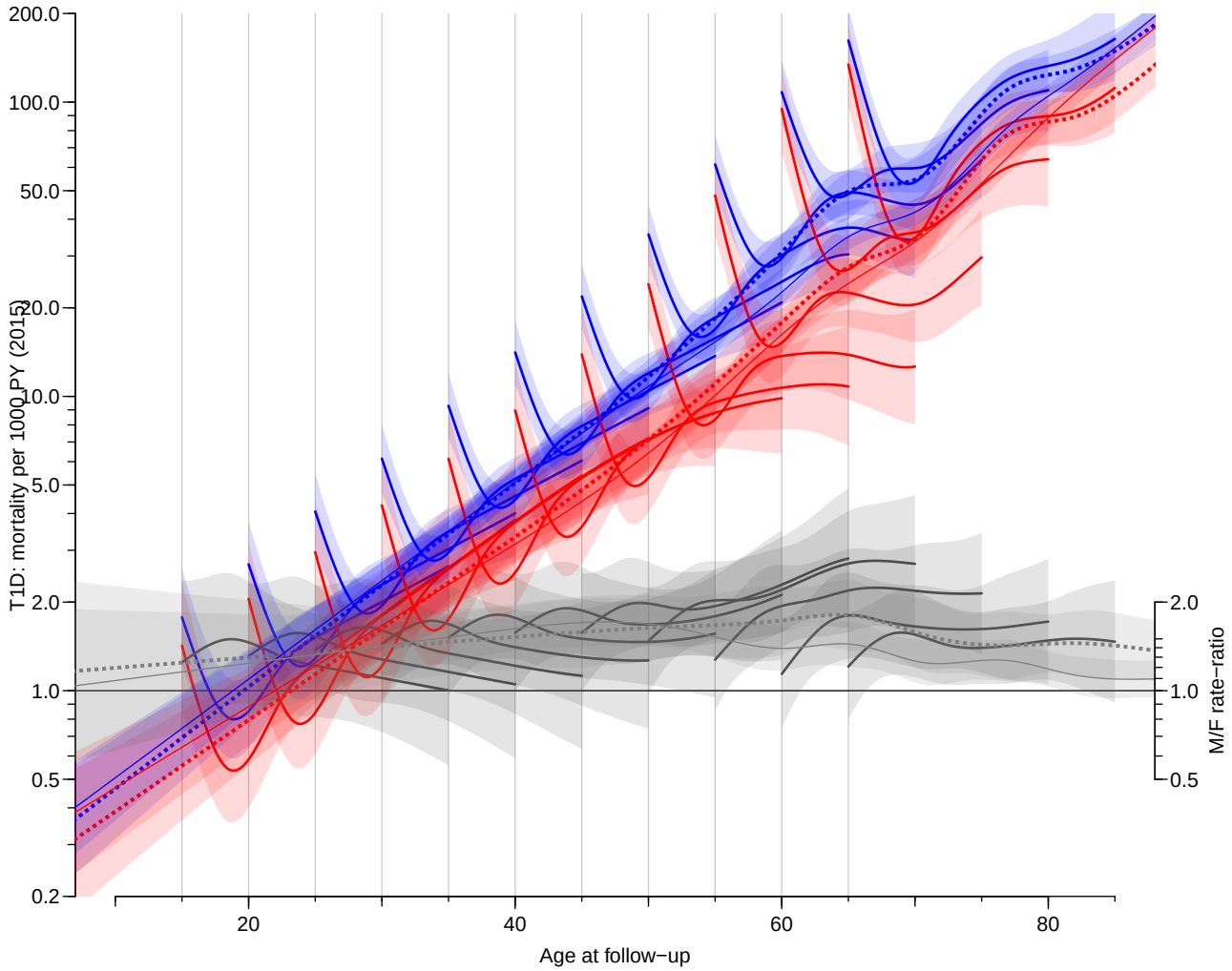


Figure 8.1: *Rates of T1D mortality by age and duration of diabetes, at 2015-01-01. Each full curve corresponds to a fixed age at diagnosis, starting at the age at diagnosis. Dotted curves are from the model where age at diagnosis and duration is ignored, the thin full lines are from including prevalent cases as of 1996-01-01 in the analysis ignoring diabetes duration. Red curves are women, blue men, black curves are the M/W rate ratios. The shaded areas indicate 95% confidence bands.* `./graph/mortx-amort-t1`

From figure 8.1 it is pretty obvious that longer diabetes duration (or younger age at diagnosis) is associated with higher mortality at a given age.

Then we plot the period-RR, extracting the values from the model object, and also showing the cohort RRs as a check of the model fit:

```
> par( mfrow=c(1,2), mar=c(3,3,1,3), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pr.P <- seq(2005,2017,0.2)
> ct.lst <- list( data.frame(Px=pr.P),
+                 data.frame(Px=2015) )
> xv <- c("Ax","dur")
> t1rr <- cbind( ci.exp( mt1, ct.lst, xvar=xv),
+                ci.exp( ft1, ct.lst, xvar=xv) )
> plot( NA, log="y", xlim=c(1996,2017), ylim=c(0.5,2),
+       xlab="Date of follow-up",
+       ylab="T1D: mortality RR relative to 2015" )
> abline( h=1 )
> axis( side=1, at=1996:2017, labels=NA, tcl=-0.3 )
> axis( side=2, at=5:15/10, labels=NA, tcl=-0.3 )
> matshade( pr.P, t1rr, lwd=2, lty=1, col=c("blue","red") )
> pr.C <- 1910:1990
> nd <- data.frame( Px=2015, Ax=2015-pr.C, prAP=1 )
> t1cr <- cbind( ci.pred( mr1, nd ),
+                ci.pred( fr1, nd ) )
> 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 )
> axis( side=1, at=seq(1910,1990,10), labels=NA, tcl=-0.3 )
> axis( side=2, at=5:15/10, labels=NA, tcl=-0.3 )
> matshade( pr.C, t1cr, lwd=2, lty=1, col=c("blue","red") )
```

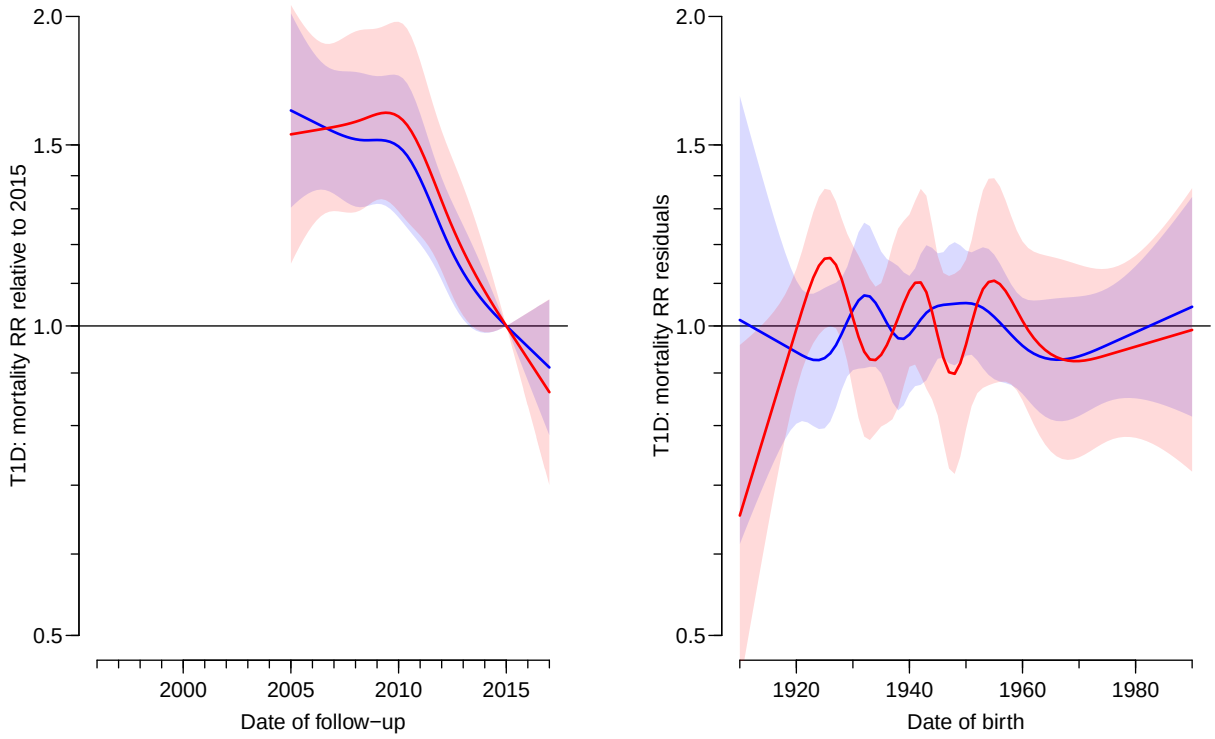


Figure 8.2: *Left panel: Rate-ratio of T1D mortality relative to 2015-01-01. Right panel: Residuals modeled by date of birth (cohort). Shaded areas indicate 95% confidence intervals; red curves are women, blue men.* `./graph/mortx-prr-crr-t1`

From figure 8.2 we se a remarkable increase in mortality from 1996 to 2010, and then an abrupt decline. The model is a model with effect of age at diagnosis, age at follow-up and duration. This may be an indication of misclassification of T1D as T2D in the period before the start of the DADD (in 2005) leaving only the survivors to be classified as T1D, conveying a smaller mortality in the earlier period.

The cohort residuals are shown as a check of model-fit, and show no indication of any systematic effect of date of birth..

8.3 Mortality for T1, T2 and all DM

Similar analyses as those above will be done for T2 patients as well as for T1 and T2 combined. Hence it will be suitable to automate it all in a single function that returns the estimated rates and RRs in a list, which then can be used to produce combined plots

```
> mort.res <-
+ function( rt1, rx1 )
+     {
+ # 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), (1:3-0.0)/4 ) ))
+ 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,
+               data = subset( rt1, sex=="M" ) )
+ ft0 <- update( mt0, data = subset( rt1, sex=="W" ) )
+ 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=="W" ) )
+ mtx <- update( mt0, data = subset( rx1, sex=="M" ) )
+ ftx <- update( mt0, data = subset( rx1, sex=="W" ) )
+ 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","W")
+ 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["W","P(lin)"] <- anova( ft1, fl1, test="Chisq" )[2,5]
+ test.tr["M","P(sl)"] <- ci.lin( ml1, subset="Px" )[,"P"]
+ test.tr["W","P(sl)"] <- ci.lin( fl1, subset="Px" )[,"P"]
+ test.tr["M",c(3,5,4)] <- (ci.exp( ml1, subset="Px" ) - 1)*100
+ test.tr["W",c(3,5,4)] <- (ci.exp( fl1, subset="Px" ) - 1)*100
```

```
+
+ # Residuals by cohort
+ prAP <- predict(mt1,type="response")
+ mr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(prAP),
+             family = poisson,
+               data = subset(rt1,sex=="M") )
+ prAP <- predict(ft1,type="response")
+ fr1 <- glm( D ~ Ns( Px-Ax, knots=c.kn ),
+             offset = log(prAP),
+             family = poisson,
+               data = subset(rt1,sex=="W") )
+
+ # Prediction data frames
+ e.pr <- seq(10,75,5)
+ d.pr <- c( NA, seq(0.1,20,0.1) )
+ 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 and M/F RRs
+ t0pr <- cbind( ci.pred( mt0, nd0 ), ci.pred( ft0, nd0 ) )
+ t0pr <- cbind( t0pr, ci.ratio(t0pr[,1:3],t0pr[,1:3+3]) )
+ txpr <- cbind( ci.pred( mtx, nd0 ), ci.pred( ftx, nd0 ) )
+ txpr <- cbind( txpr, ci.ratio(txpr[,1:3],txpr[,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
+ ct.lst <- list( data.frame(Px=seq(2005,2017,0.2)),
+                 data.frame(Px=2015) )
+ xv <- c("Ax","dur")
+ t1rr <- cbind( ci.exp( mt1, ct.lst, xvar=xv),
+                ci.exp( ft1, ct.lst, xvar=xv) )
+
+ # Cohort residuals
+ ndc <- data.frame( Px=2015, Ax=2015-(1910:1990), prAP=1 )
+ t1cr <- cbind( ci.pred( mr1, ndc ),
+                ci.pred( fr1, ndc ) )
+
+ # return results
+ list( tr = test.tr,
+       A0 = data.frame(nd0[,"Ax"],t0pr,txpr),
+       Ad = data.frame(nd[,c("Ae","dur","Ax")],t1pr),
+       Pr = data.frame(P=pr.P,t1rr),
+       Cr = data.frame(C=pr.C,t1cr) )
+ }
> # Now use the function to generate results for T1, T2 and combined DM
> rT1 <- mort.res( subset( rtL, state=="T1" ),
+                  subset( rt , state=="T1" ) )
> rT2 <- mort.res( subset( rtL, state=="T2" ),
+                  subset( rt , state=="T2" ) )
> rDM <- mort.res( subset( rtL, state %in% c("T1","T2") ),
+                  subset( rt , state %in% c("T1","T2") ) )
```

```
> save( rT1, rT2, rDM, file="../nydata/mortx-res.Rda" )
```

An overview of the annual changes in mortality:

```
> load( file="../nydata/mortx-res.Rda" )
> fCp( rT1$tr, d=3, z=NULL )
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.208     0.000    -5.095    -3.534    -6.631
W     0.072     0.000    -4.977    -2.878    -7.030
> fCp( rT2$tr, d=3, z=NULL )
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.019     0.000    -3.640    -3.359    -3.920
W     0.000     0.000    -3.063    -2.742    -3.383
> fCp( rDM$tr, d=3, z=NULL )
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.009     0.000    -3.872    -3.597    -4.147
W     0.000     0.000    -3.239    -2.923    -3.554
```

8.4 Plotting mortality rates

We will make two sets of plots with 2 rows indexed by T1 / T2; one set with mortality, period RR and cohort residuals aimed for the ESM and one with only mortality rates and period RR aimed at the paper itself.

```
> pltp <-
+ function( rT1, coh=FALSE, e.pr=seq(10,70,10),
+           lab="", rlim=c(0.2,200), rrpos=1, nlt=NULL, rr=FALSE, article=TRUE )
+     {
+ # utilities for letter-labeling
+ ulc <- function(){uu<-par("usr")
+        list(x=uu[1]*0.93+uu[2]*0.07, y=10^(uu[4]*0.97+uu[3]*0.03)) }
+ nxl <- function(){nlt<<-nlt+1
+        text(ulc(),paste(letters[nlt],lab,sep="   "),font=1,cex=1.3,adj=0) }
+
+ rdf <- subset( rT1$Ad, Ae %in% e.pr )
+ art <- rT1$A0
+ prt <- rT1$Pr
+ crt <- rT1$Cr
+
+ # Adjust the RRs
+ rdf[,10:12] <- rdf[,10:12] * rrpos
+ art[,c(8:10,17:19)] <- art[,c(8:10,17:19)] * rrpos
+
+ # nice plotting frame for rates
+ plot( NA, log="y",
+       xlim=c(10,85), ylim=rlim, yaxs="i", xlab="", ylab="" )
+ abline( v=e.pr, col=gray(0.8) )
+ abline( h=rrpos )
+ if( rr ) abline( h=1 )
+ # axis grooming
+ axis( side=1, at=seq(10,85,10), labels=NA, tcl=-0.4 )
+ axis( side=1, at=seq(10,85, 5), labels=NA, tcl=-0.2 )
+ axis( side=2, at=outer(2:10,-1:2,function(x,y) x*10^y), labels=NA, tcl=-0.3 )
```

```
+ axis( side=4, at=c(5,10,20)/10*rrpos, labels=c(5,10,20)/10 )
+ axis( side=4, at=c(5:9,15,20)/10*rrpos, labels=NA, tcl=-0.3 )
+ mtext( "M/W\nRR", side=4, at=rrpos*3, line=0 )
+ # rates plotted
+ for( e in e.pr )
+ matshade( rdf$Ax[rdf$Ae==e], rdf[rdf$Ae==e,-(1:3)],
+           type="l", lwd=2, lty=1,
+           col=c("blue","red",gray(0.3)), )
+ if( article ) {
+ # matshade( art[,1], art[,2:10],
+ matlines( art[,1], art[,c(2,5,8)],
+           type="l", lwd=3, lty="11", lend="butt",
+           col=c("blue","red",gray(0.5)) )
+ # matshade( art[,1], art[,11:19],
+ matlines( art[,1], art[,c(11,14,17)],
+           type="l", lwd=1, lty=1, lend="butt",
+           col=c("blue","red",gray(0.5)) )
+           }
+ if( !is.null(nlt) ) nxl()
+
+ plot( NA, log="y", xlim=c(1996,2017), ylim=c(0.4,2.5),
+       xlab="", ylab="" )
+ abline( h=1 )
+ axis( side=1, at=1996:2017, labels=NA, tcl=-0.3 )
+ axis( side=2, at=4:15/10, labels=NA, tcl=-0.3 )
+ matshade( prt$P, prt[,-1], lwd=2, lty=1, col=c("blue","red") )
+ if( !is.null(nlt) ) nxl()
+
+ if( coh ){
+ plot( NA, log="y", xlim=c(1910,1990), ylim=c(0.4,2.5),
+       xlab="", ylab="" )
+ abline( h=1 )
+ axis( side=1, at=seq(1910,1990,10), labels=NA, tcl=-0.3 )
+ axis( side=2, at=4:15/10, labels=NA, tcl=-0.3 )
+ matshade( crt$C, crt[,-1], lwd=2, lty=1, col=c("blue","red") )
+ if( !is.null(nlt) ) nxl()
+          }
+     }
```

With this function defined we plot a 2 by 3 layout aimed at the ESM:

```
> layout( matrix(1:6,2,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rT1, coh=TRUE, nlt=0, e.pr=1:5*10+5 )
> pltp( rT2, coh=TRUE, nlt=3, e.pr=3:7*10 )
> mtext( c("Mortality rate per 1000 PY",
+          "RR relative to 2015",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE, cex=0.67 )
> 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(1:2,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(1,1,1,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rT1, coh=FALSE, lab="T1D", nlt=0, e.pr=1:5*10+5,
+       rlim=c(0.1,200), rrpos=0.2, article=FALSE )
> mtext( c("Mortality rate per 1000 PY",
+          "Rate ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )


> layout( matrix(1:2,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(1,1,1,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rT2, coh=FALSE, lab="T2D", nlt=0, e.pr=3:7*10,
+       rlim=c(0.1,200), rrpos=0.2, article=FALSE )
> mtext( c("Mortality rate per 1000 PY",
+          "Rate ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

For the paper/slide proper we leave out the cohort residuals:

```
> layout( matrix(1:4,2,2,byrow=T), widths=c(8,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rT1, coh=FALSE, nlt=0, e.pr=1:5*10+5, rlim=c(0.1,200), rrpos=0.2 )
> pltp( rT2, coh=FALSE, nlt=2, e.pr=3:7*10, rlim=c(0.1,200), rrpos=0.2 )
> mtext( c("Mortality rate per 1000 PY",
+          "Rate ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

Finally, we make the same set of plots, but for all diabetes combined (it is going to be quite close to the T2D results)

```
> layout( matrix(1:3,1,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( rDM, coh=TRUE, nlt=0, e.pr=c(15,30,45,60,75) )
> mtext( c("Mortality rate per 1000 PY",
+          "RR relative to 2015",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE, cex=0.67 )
> 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 )
```

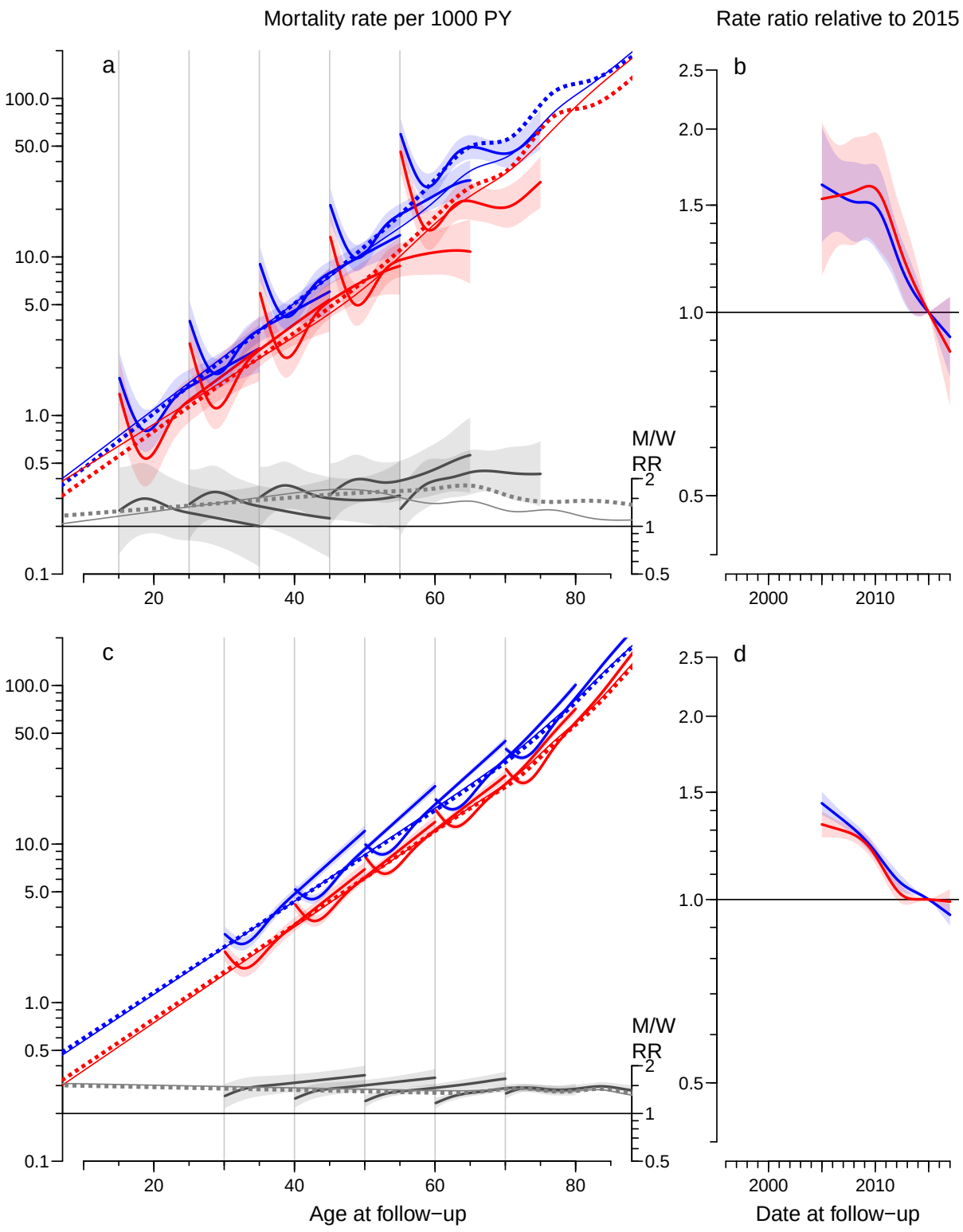


Figure 8.3: *Mortality and mortality RR relative to 2015-01-01. Left panels show the mortality rates at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are mortality rates modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall mortality also including prevalent cases as of 1996-01-01. The right panels show the mortality RR relative to 2015-01-01.*
Red curves are for women, blue for men, black are M/W RR; shaded areas indicate 95% confidence intervals. `./graph/mortx-art-m`

`mortx.tex`

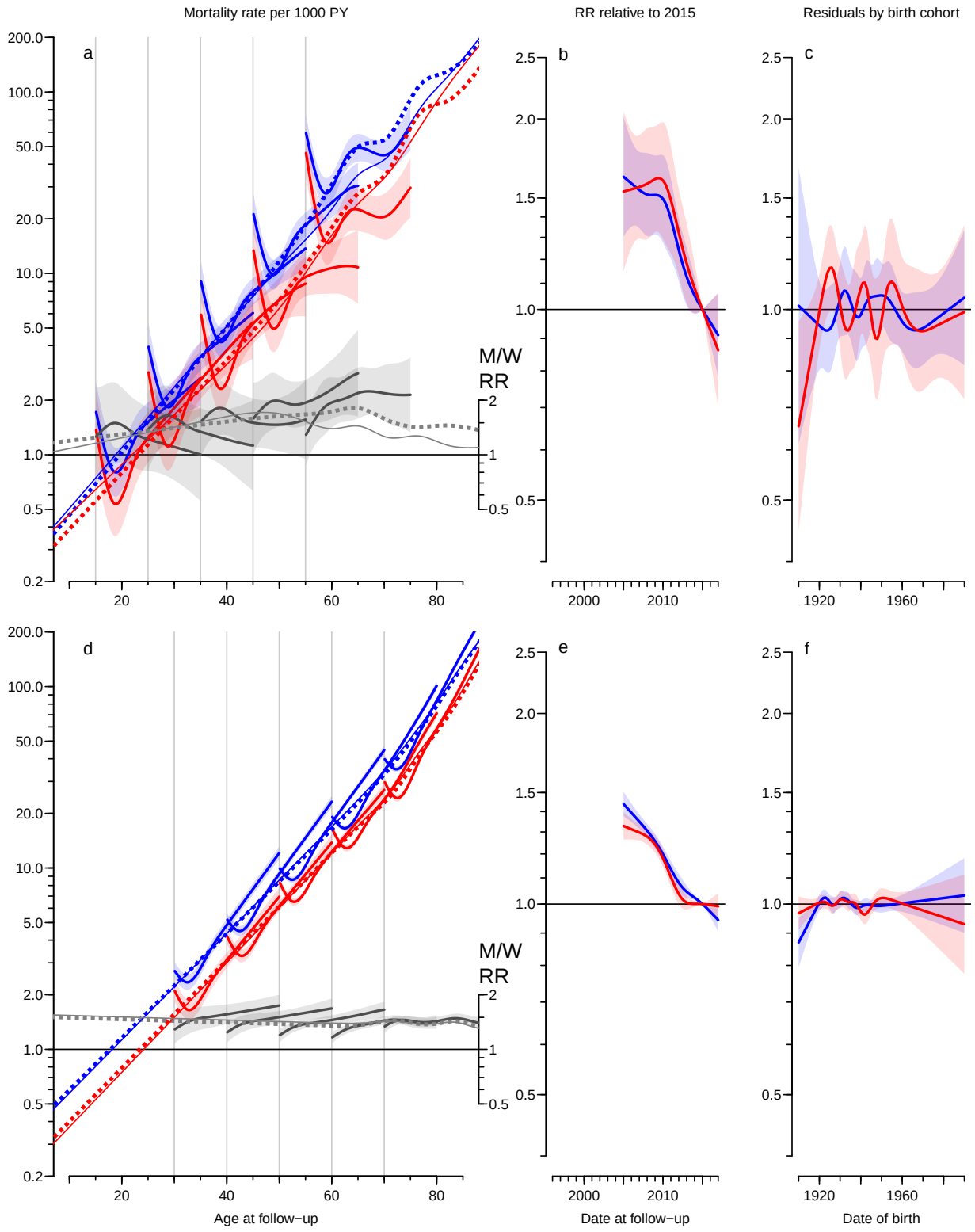


Figure 8.4: *Mortality, and RR relative to 2015-01-01 and birth cohort residuals. Leftmost plot shows the mortality rates at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are mortality rates modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall mortality also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W RR; shaded areas indicate 95% confidence intervals. `./graph/mortx-all-m`
`mortx.tex`

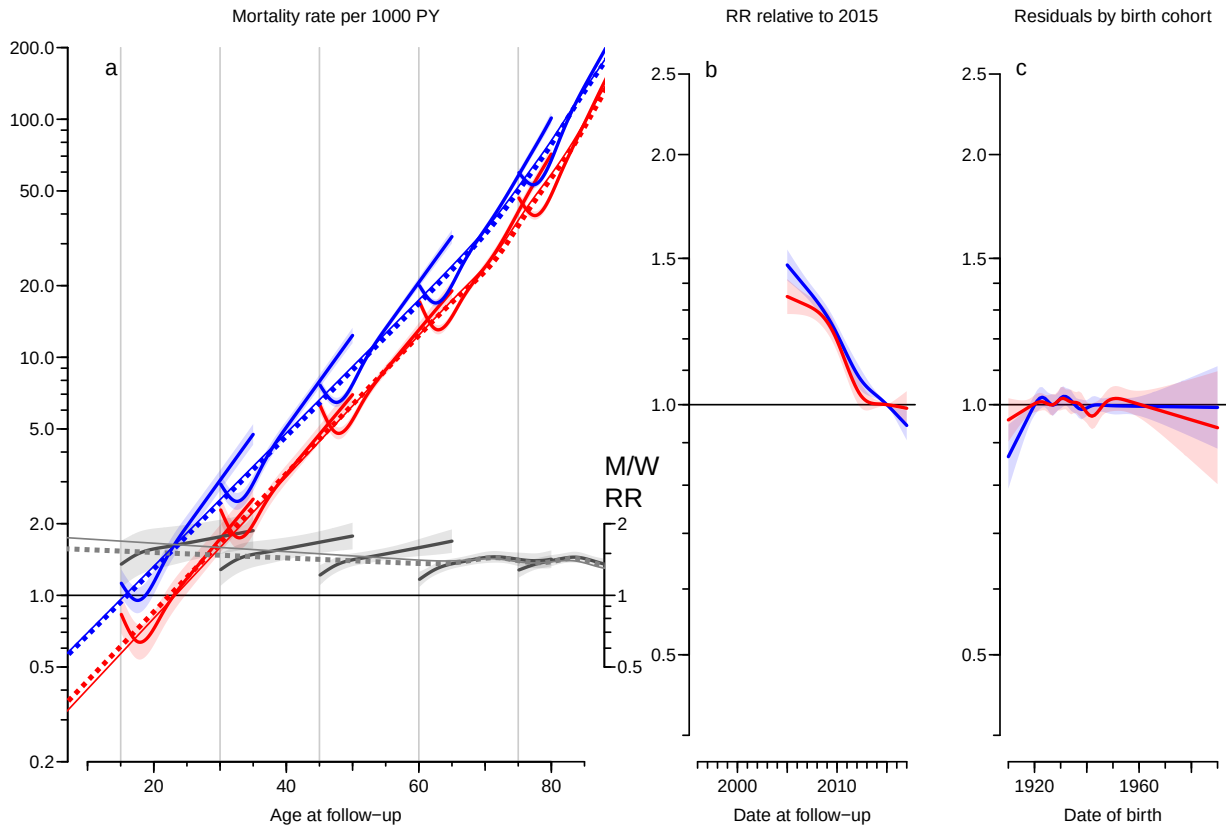


Figure 8.5: *Mortality, HR relative to 2015-01-01 and birth cohort residuals. Left panel shows the mortality rates at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken line in left panel is mortality rates modeled ignoring age at diagnosis and duration of diabetes. Thin full line is overall mortality also including prevalent cases as of 1996-01-01. Right panel is the mortality RR relative to 2015-10-01. Red curves are for women, blue for men, black are M/W RR; shaded areas indicate 95% confidence intervals.* `./graph/mortx-DM-m`

For the paper proper we leave out the cohort residuals:

```
> layout( matrix(1:4,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n", cex=0.84 )
> pltp( rDM, coh=FALSE, nlt=0, e.pr=c(15,30,45,60,75), rlim=c(0.1,200), rrpos=0.2 )
> mtext( c("Mortality rate per 1000 PY",
+          "RR relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

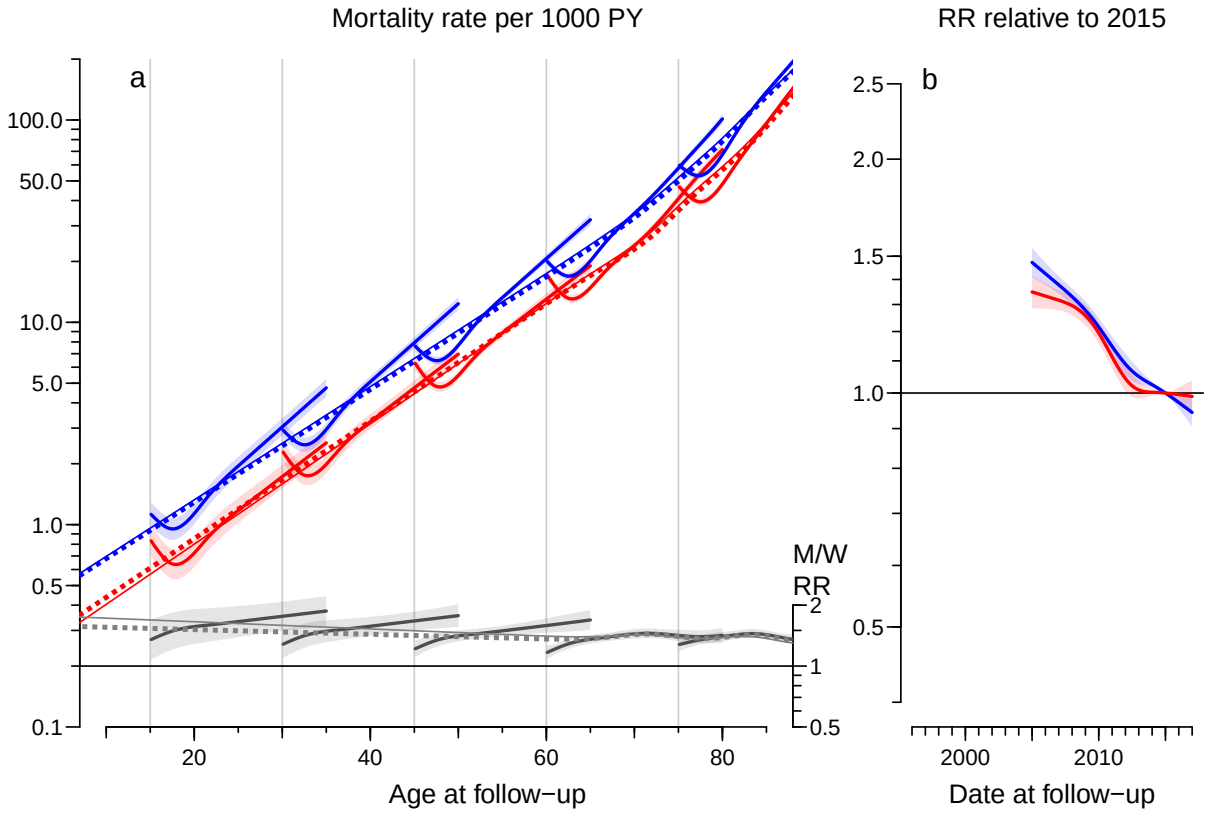


Figure 8.6: *Mortality and mortality RR relative to 2015-01-01. Leftmost plot shows the mortality rates at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are mortality rates modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall mortality also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W RR; shaded areas indicate 95% confidence intervals. `./graph/mortx-aDM-m`

8.5 Comparison of T1 and T2 mortality

Since we fitted models separately for T1 and T2 diabetes we can — just as we did for the M/W comparison — derive mortality rate-rations between T1 and T2 mortality.

Specifically, the models we reported from were for current age, a, age at diagnosis e and diabetes duration $d = a - e$ at calendar time p:

$$\log\big(\lambda(a,e,d,p)\big) = f(a) + g(e) + h(d) + j(p)$$

This was fitted separately for each sex and diabetes type, so for a given sex, the T2/T1 mortality rate-ratio is:

$$\log\big(\mathrm{RR}(a,e,d,p)\big) = \big(f_2(a) - f_1(a) + g_2(e) - g_1(e) + h_2(d) - h_1(d)\big) + \big(j_2(p) - j_1(p)\big)$$

So we make a graph similar to the previous ones, but now for the T2/T1 ratios, using each of the two terms in the larger brackets.

We can even produce the M/W ratios of mortality RRs, exploring whether the two types of diabetes have the same impact on mortality for men and women.

Since calculations are done on separate subsets of data, calculation of confidence intervals are straight forward. The 12 variables in `rT1$Ad` are the age/diag/dur terms from the mortality models for T1, the first three variables are age at entry (`Ae`), diabetes duration `dur` and age at follow-up (`Ax`), and the remaining 9 columns are the mortality rates for men, women and the M/W-RR of rates. Thus all we need is to take the ratio of these between T1 and T2 (computing the confidence limits too, of course).

Similar calculations can be done for the structures `A0` ignoring the duration and the `Px` of the secular trends.

```
> RR0 <- rT1$A0
> RRd <- rT1$Ad
> RRp <- rT1$Pr
> for( i in 0:5 ) {wh<-1+i*3+(1:3) ; RR0[,wh] <- ci.ratio( as.matrix(rT1$A0[,wh]),
+                                                          as.matrix(rT2$A0[,wh]) ) }
> for( i in 0:2 ) {wh<-3+i*3+(1:3) ; RRd[,wh] <- ci.ratio( as.matrix(rT1$Ad[,wh]),
+                                                          as.matrix(rT2$Ad[,wh]) ) }
> for( i in 0:1 ) {wh<-1+i*3+(1:3) ; RRp[,wh] <- ci.ratio( as.matrix(rT1$Pr[,wh]),
+                                                          as.matrix(rT2$Pr[,wh]) ) }
```

Finally we put structures in a list so it can be referenced in the function `pltp`:

```
> RR <- list( A0=RR0, Ad=RRd, Pr=RRp )
```

We simply re-use the code for the mortality curves and plot the graphs not of mortality rates, but of T1D vs. T2D RRs:

```
> layout( matrix(1:4,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n", cex=0.83 )
> pltp( RR, coh=FALSE, nlt=0, rlim=c(0.1,5), e.pr=c(30,45,60), rrpos=0.2, rr=TRUE )
> # For illustration a decline of 3\%/year
> x <- 2010:2017 ; y <- 0.97^(x-2015) ; lines(x,y,lty=3)
> mtext( c("T1D vs. T2D mortality rate-ratio at 2015",
+          "RR ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date of follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

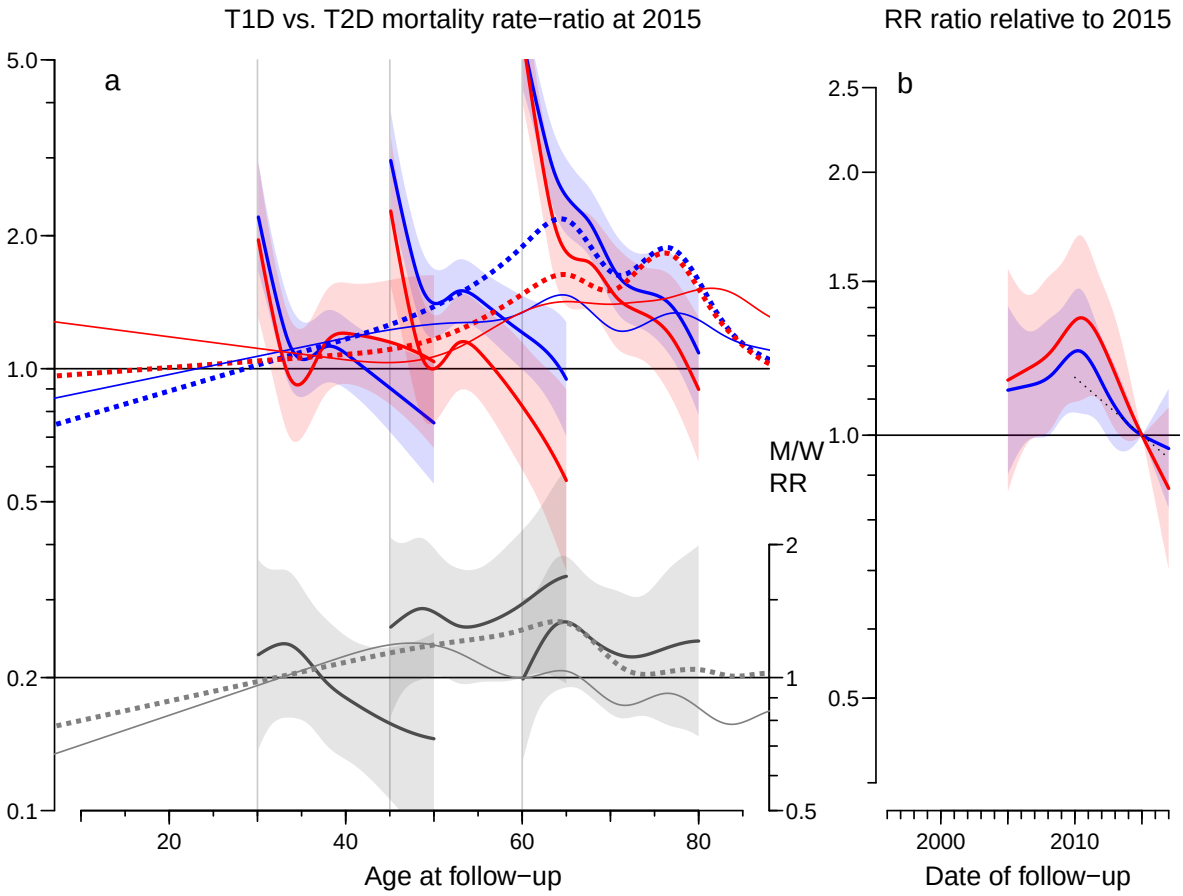


Figure 8.7: *T1D versus T2D mortality RR at 2015-01-01. Leftmost plot shows the mortality RR at 2015-01-01 for persons diagnosed in different ages.*
Broken lines in leftmost plot are mortality RRs modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall mortality RR also including prevalent cases as of 1996-01-01.
Red curves are for women, blue for men, black are M/W RR ratio; shaded areas indicate 95% confidence intervals. `./graph/mortx-t1t2-rr`

From figure 8.7 we see that in the range of age at diagnosis 40–60, the T1D vs T2D mortality rate ratio is broadly 2 for women and 2.5 for men during the first few years after diagnosis, and after that attenuates by duration / age. Ignoring duration of diabetes and age at diagnosis, showed a T1/T2 rate ratio of 1.5 for man and 1.2 for women at age 50. In the range 30–70 years of age the T1/T2 mortality RR was higher among men than among women; outside this interval there was essentially no detectable difference in RRs between man and women.

The trend in RR in the right panel shows that the RR among T1D patients have been declining by some 3% per year since 2010 (direct measurement on the graph, the broken black line has a slope of -3%/year).

8.5.1 Direct comparison of T1D vs. T2D mortality

We can derive anm overall mortality RR between T1D and T2D by fitting a joint model for T1D and T2D pateinst taking current age, duration, at at diagnosis and period into account:

```
> rtD <-  subset( rtL, state!="noDM" )
> a.kn <-     with( rtD, quantile( rep(Ax    ,D), (1:8-0.5)/8 ) )
> d.kn <- c(0,with( rtD, quantile( rep(   dur,D), (1:3-0.0)/4 ) ))
> e.kn <-     with( rtD, quantile( rep(Ax-dur,D), (1:5-0.5)/5 ) )
> p.kn <-     with( rtD, quantile( rep(   Px ,D), (1:5-0.5)/5 ) )
> c.kn <-     with( rtD, quantile( rep(Px-Ax ,D), (1:9-0.5)/9 ) )
> b12 <- glm( D ~ Ns( Ax    , knots=a.kn ) +
+                 Ns(    dur, knots=d.kn ) +
+                 Ns( Ax-dur, knots=e.kn ) +
+                 Ns( Px    , knots=p.kn ) +
+                 Ns( Cx    , knots=c.kn ) + state + sex,
+             offset = log(Y),
+             family = poisson,
+               data = rtD )
> m12 <- update( b12, . ~ . - sex, data = subset( rtD, sex=="M" ) )
> w12 <- update( b12, . ~ . - sex, data = subset( rtD, sex=="W" ) )
> round( cbind(
+        ci.exp( m12, subset="st", ctr.mat=rbind( 1, -1) ),
+        ci.exp( w12, subset="st", ctr.mat=rbind( 1, -1) ),
+        ci.exp( b12, subset="st", ctr.mat=rbind( 1, -1) ) ), 2 )
     exp(Est.) 2.5% 97.5% exp(Est.) 2.5% 97.5% exp(Est.) 2.5% 97.5%
[1,]      0.53 0.51  0.56      0.59 0.55  0.63      0.55 0.53  0.58
[2,]      1.87 1.77  1.97      1.71 1.59  1.83      1.81 1.73  1.88
```

Thus we see that there is an overall different T1 / T2 mortality rate ratio between man and women, for man it is 1.86, for women 1.55. This is controlled for age at FU, duration, age at diagnosis, calendar time at FU, and date of birth.

8.6 SMR

The analysis of SMR is parallel to the analysis of rates, except that we replace the person-years by the expected number of events. This is however easily done by creating a new dataset with expected numbers; first by extraction the rates (per 1000 PY) from the noDM part of data:

```
> names(rt)
```

```
 [1] "sex"   "state" "A"     "P"     "C"     "dur"   "Dcvd"  "Dcan"  "Dres"  "Doth"  "Y"
[12] "T1"    "T2"    "D"     "Ax"    "Px"    "Cx"
> nDM <- transform( subset( rt,
+                           state=="noDM",
+                           select=c("sex","Ax","Px","D","Y") ),
+                   R = D / Y )[,c("sex","Ax","Px","R")]
> table( with( nDM, table(sex,Ax,Px) ) )
   0    1
4800 4800
> table( with( nDM, table(sex,floor(Ax),Px) ) )
   1
4800
> table( with( nDM, table(sex,floor(Ax),floor(Px)) ) )
   2
2400
```

— then merge with the original data computing the expected numbers. However, first note that it is immaterial what units `Y` is measured in, as long as it is in the same units across the dataset `rt`. Second, note that we are naming the variable with expected numbers "`Y`", so we can use the same code as before:

```
> xt  <- transform( merge( rt , nDM, all=TRUE ),
+                   Y = Y * R )
> xtL <- transform( merge( rtL, nDM, all=TRUE ),
+                   Y = Y * R )
> # check that Obs=Exp in the
> round( do.call( rbind, with( xt, tapply( D-Y, state, range ) ) ), 4 )
        [,1]    [,2]
noDM  0.0000  0.0000
T1   -0.6284  4.7229
T2   -4.5926 12.9185
```

Now use the `mort.res` function to generate SMR results for T1, T2 and combined DM:

```
> xT1 <- mort.res( subset( xtL, state=="T1" & Y>0 ),
+                  subset( xt , state=="T1" & Y>0 ) )
> xT2 <- mort.res( subset( xtL, state=="T2" & Y>0 ),
+                  subset( xt , state=="T2" & Y>0 ) )
> xDM <- mort.res( subset( xtL, state %in% c("T1","T2") & Y>0 ),
+                  subset( xt , state %in% c("T1","T2") & Y>0 ) )
> save( xT1, xT2, xDM, file="../nydata/smrx-res.Rda" )
```

An overview of the annual changes in SMR:

```
> load( file="../nydata/smrx-res.Rda" )
> fCp( xT1$tr, d=3, z=NULL )
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.380     0.022    -1.891    -0.275    -3.481
W     0.133     0.043    -2.237    -0.070    -4.358
> fCp( xT2$tr, d=3, z=NULL )
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.107     0.000    -0.722    -0.432    -1.010
W     0.000     0.000    -0.595    -0.265    -0.923
> fCp( xDM$tr, d=3, z=NULL )
  P(lin)    P(sl)     Chg %/y   lo        hi
M     0.096     0.000    -0.951    -0.667    -1.234
W     0.000     0.000    -0.769    -0.445    -1.093
```

8.6.1 Plotting the SMR

With the function `pltp` we plot a 2 by 3 layout aimed at the ESM:

```
> layout( matrix(1:6,2,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,1), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xT1, coh=TRUE, nlt=0, rlim=c(0.3,20), e.pr=1:5*10+5 )
> pltp( xT2, coh=TRUE, nlt=3, rlim=c(0.3,20), e.pr=3:7*10 )
> mtext( c("SMR",
+          "SMR ratio relative to 2015",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE, cex=0.67 )
> 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 )
```

For the paper proper we leave out the cohort residuals:

```
> layout( matrix(1:4,2,2,byrow=T), widths=c(8,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xT1, coh=FALSE, nlt=0, rlim=c(0.3,20), e.pr=1:5*10+5 )
> pltp( xT2, coh=FALSE, nlt=2, rlim=c(0.3,20), e.pr=3:7*10 )
> mtext( c("SMR at 2015",
+          "SMR ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE, cex=0.66 )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE, cex=0.66 )
```

. . . and separate versions for slides:

```
> layout( matrix(1:2,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(1,1,2,0), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xT1, coh=FALSE, lab="T1D", nlt=0, rlim=c(0.3,20), e.pr=1:5*10+5 )
> mtext( c("SMR at 2015",
+          "SMR ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )


> layout( matrix(1:2,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(1,1,2,0), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xT2, coh=FALSE, lab="T2D", nlt=2, rlim=c(0.3,20), e.pr=3:7*10 )
> mtext( c("SMR at 2015",
+          "SMR ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

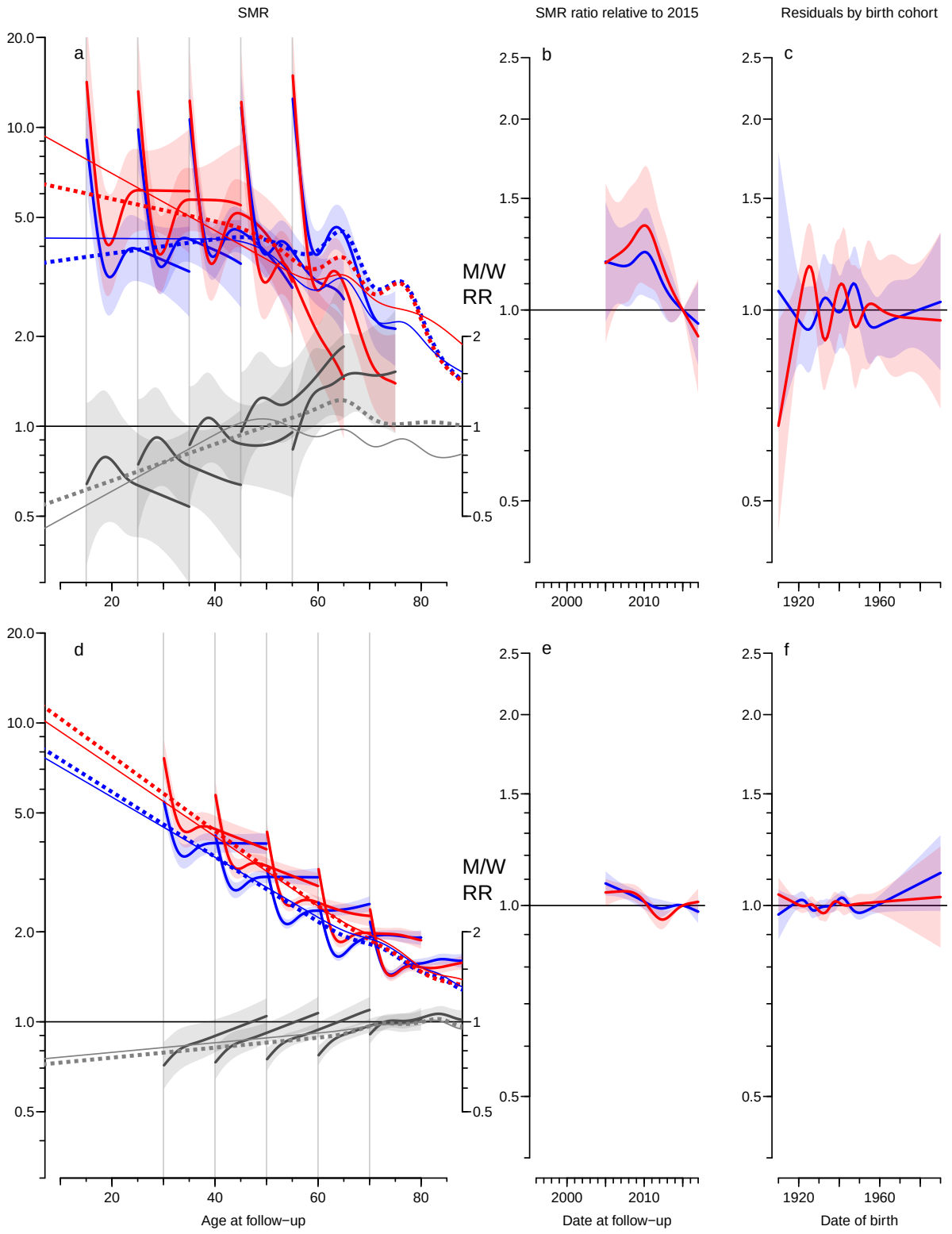


Figure 8.8: *SMR and SMR-ratio relative to 2015-01-01 and birth cohort residuals. Leftmost plot shows the SMR at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are SMR modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall SMR also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W SMR ratios; shaded areas indicate 95% confidence intervals. `./graph/mortx-all-smr`
`mortx.tex`

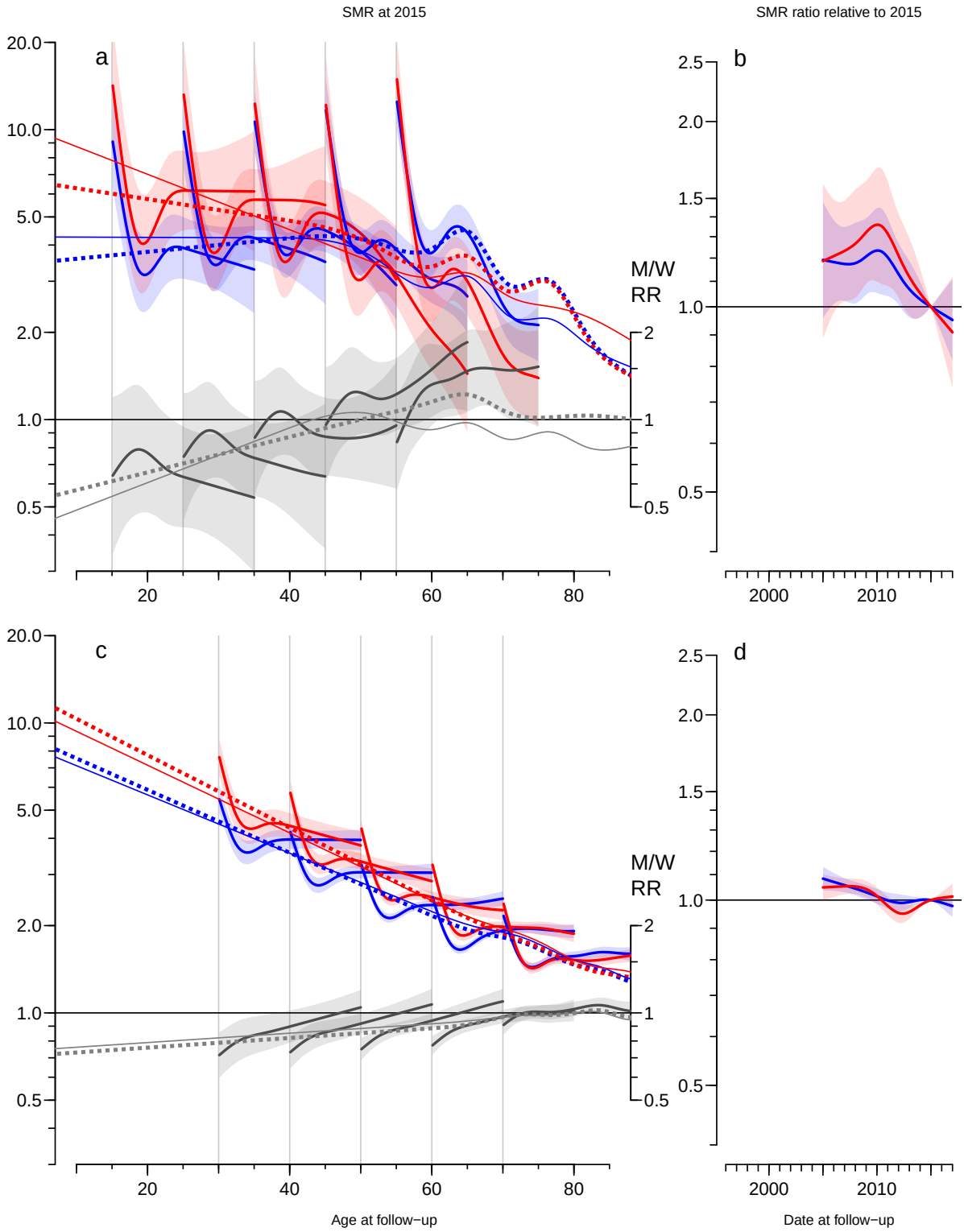


Figure 8.9: *SMR and SMR-ratio relative to 2015-01-01. Leftmost plot shows the SMR at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are SMR modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall SMR also including prevalent cases as of 1996-01-01. Red curves are for women, blue for men, black are M/W SMR ratios; shaded areas indicate 95% confidence intervals.* `./graph/mortx-art-smr`

Finally we make the same set of plots, but for all diabetes combined (it is going to be quite close to the T2D results)

```
> layout( matrix(1:3,1,3,byrow=T), widths=c(8,4,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xDM, coh=TRUE, lab="DM", nlt=0, rlim=c(0.3,30), e.pr=c(15,30,45,60,75) )
> mtext( c("SMR",
+          "SMR ratio relative to 2015",
+          "Residuals by birth cohort"),
+        at = c(4,10,14)/16,
+        side=3, outer=TRUE, cex=0.84 )
> mtext( c("Age at follow-up",
+          "Date at follow-up",
+          "Date of birth"),
+        at = c(4,10,14)/16,
+        side=1, outer=TRUE )
```

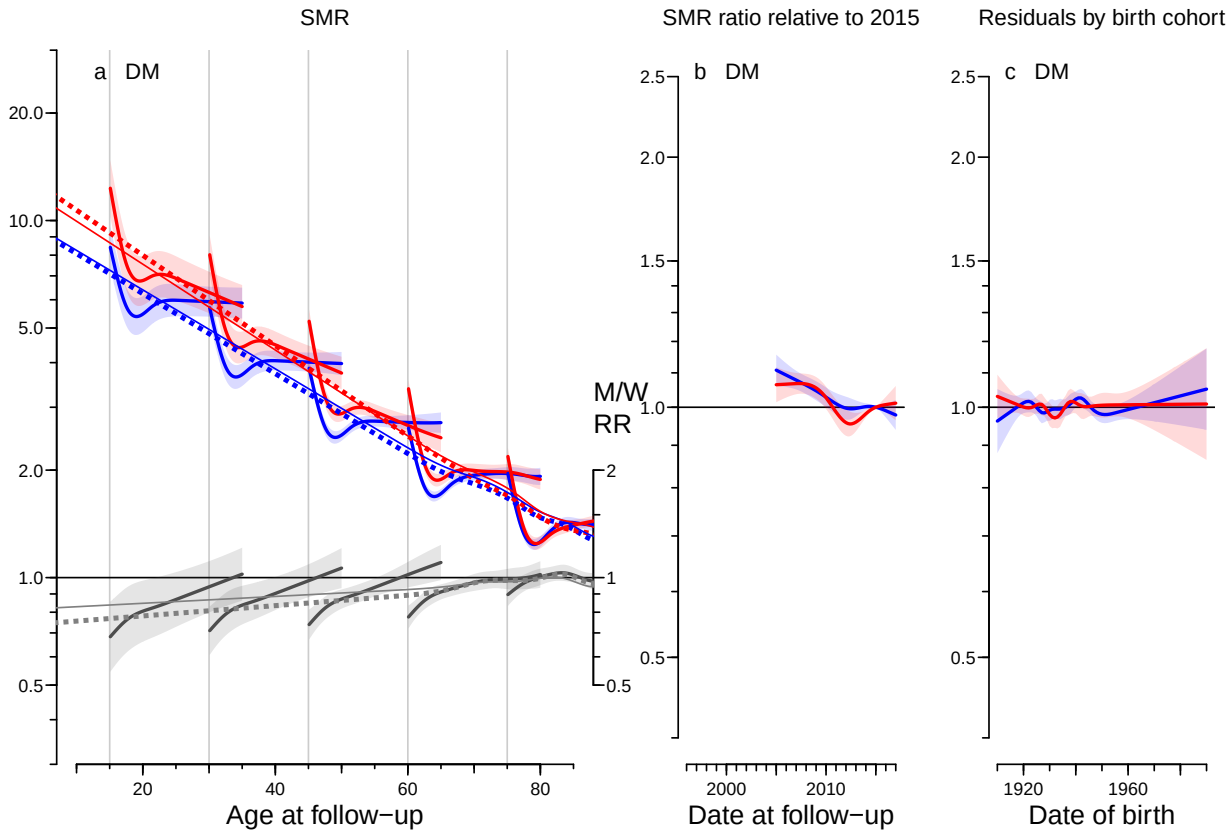


Figure 8.10: *SMR and SMR-ratio for all DM, relative to 2015-01-01 and birth cohort residuals. Leftmost plot shows the SMR at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are SMR modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall SMR also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W SMR ratios; shaded areas indicate 95% confidence intervals. `./graph/mortx-DM-smr`

For the paper proper we leave out the cohort residuals:

```
> layout( matrix(1:4,1,2,byrow=T), widths=c(8,4) )
> par( oma=c(3,2,2,0), mar=c(2,2,1,2), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pltp( xDM, coh=FALSE, lab="DM", nlt=0, rlim=c(0.3,20), e.pr=c(15,30,45,60,75) )
> mtext( c("SMR",
+          "SMR ratio relative to 2015"),
+        at = c(4.5,10)/12,
+        side=3, outer=TRUE )
> mtext( c("Age at follow-up",
+          "Date at follow-up"),
+        at = c(4.5,10)/12,
+        side=1, outer=TRUE )
```

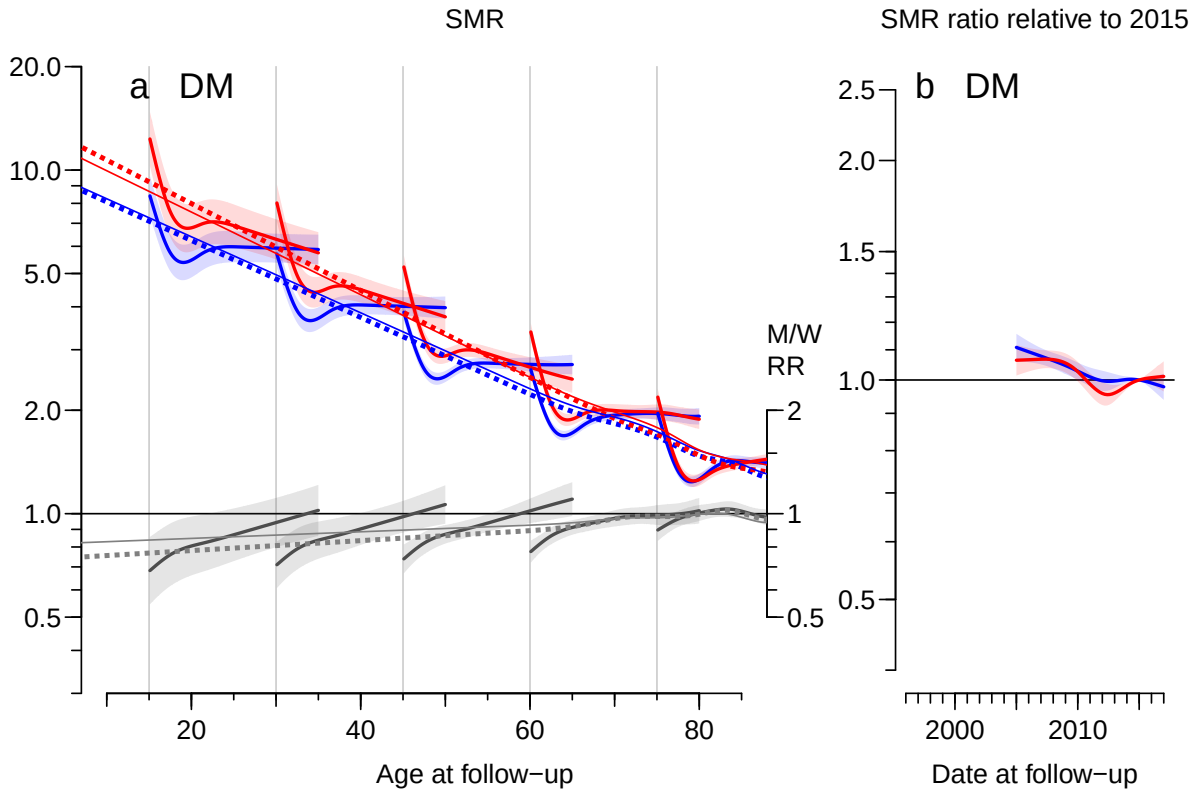


Figure 8.11: *SMR and SMR-ratio for all DM, relative to 2015-01-01. Leftmost plot shows the SMR at 2015-01-01 for persons diagnosed in ages 10, 20, . . . , 70, followed for 0–20 years of diabetes duration. Broken lines in leftmost plot are SMR modeled ignoring age at diagnosis and duration of diabetes. Thin full lines are overall SMR also including prevalent cases as of 1996-01-01.*
Red curves are for women, blue for men, black are M/W SMR ratios; shaded areas indicate 95% confidence intervals. `./graph/mortx-aDM-smr`

```
----------------------
2019-01-10 at 18:10:59
Time elapsed: 00:00:58
----------------------
```

. . . now input from `pr-rates.tex`

Chapter 9

Analysis and prediction of rates

```
> library( Epi )
> library( splines )
> library( mgcv )
> start(FALSE)
> load( file="../nydata/inits.Rda" )
> load( file="../nydata/rt.Rda" )
```

9.1 Analysis data

We model the incidence and mortality rates from a tabulation of the diabetes register made in conjunction with the annual update of the register. The analytical units are the Lexis triangles ($\triangledown$ and $\triangle$), classified by age, period and cohort:

```
> head( rt )
  sex state A    P    C  dur Dcvd Dcan Dres Doth            Y T1 T2 D        Ax       Px
1   M    T1 0 1997 1996 0.10    0    0    0    0 1.430527e-04  0  0 0 0.6666667 1997.333
2   M    T1 0 1999 1998 0.10    0    0    0    0 2.258727e-05  0  0 0 0.6666667 1999.333
3   M    T1 0 1999 1999 0.10    0    0    0    0 7.665982e-05  0  0 0 0.3333333 1999.667
4   M    T1 0 2000 1999 0.10    0    0    0    0 3.793292e-04  0  0 0 0.6666667 2000.333
5   M    T1 0 2000 1999 0.35    0    0    0    0 1.757700e-04  0  0 0 0.6666667 2000.333
6   M    T1 0 2001 2000 0.10    0    0    0    0 7.734428e-05  0  0 0 0.6666667 2001.333
        Cx
1 1996.667
2 1998.667
3 1999.333
4 1999.667
5 1999.667
6 2000.667
> table( rt$Y>0 )
  TRUE
354350
> range( rt$Ax )
[1]  0.3333333 99.6666667
> range( rt$Px )
[1] 1996.333 2016.667
```

— modeling of rates requires that the person-years are strictly positive.

A brief overview of the number of events and PY:

```
> tt <- xtabs( cbind(T1,T2,DM=T1+T2,D,Y=Y*1000) ~ state + P,
+              data = rt )
> fCtable( addmargins(tt,1:2), row.vars=c(1,2), w=11 )
                    T1          T2          DM           D           Y
state P
noDM  1996       1,194      11,405      12,599      54,263   5,193,867
      1997       1,173      10,757      11,930      53,448   5,208,856
      1998       1,111      11,824      12,935      51,957   5,220,853
      1999       1,005      12,446      13,451      52,394   5,230,170
      2000         988      12,197      13,185      50,642   5,241,289
      2001       1,001      12,244      13,245      51,174   5,254,589
      2002         988      15,356      16,344      50,935   5,265,177
      2003         931      16,819      17,750      49,744   5,269,785
      2004         897      17,010      17,907      47,752   5,273,131
      2005         896      14,684      15,580      46,829   5,278,757
      2006         936      14,112      15,048      46,616   5,289,351
      2007         948      15,530      16,478      47,020   5,303,997
      2008         913      17,400      18,313      45,666   5,323,692
      2009         925      18,482      19,407      45,585   5,341,701
      2010         896      20,571      21,467      44,658   5,354,176
      2011         854      28,743      29,597      42,860   5,361,742
      2012         801      22,799      23,600      42,099   5,365,666
      2013         822      18,186      19,008      41,747   5,378,276
      2014         806      17,241      18,047      40,476   5,402,571
      2015         851      17,625      18,476      41,514   5,437,983
      2016         776      18,521      19,297      41,190   5,460,435
      Sum       19,712     343,952     363,664     988,569 111,456,064
T1    1996           .           .           .         616      22,220
      1997           .           .           .         664      22,736
      1998           .           .           .         686      23,145
      1999           .           .           .         706      23,458
      2000           .           .           .         721      23,739
      2001           .           .           .         708      23,967
      2002           .           .           .         750      24,214
      2003           .           .           .         759      24,352
      2004           .           .           .         694      24,515
      2005           .           .           .         711      24,645
      2006           .           .           .         688      24,836
      2007           .           .           .         695      24,898
      2008           .           .           .         639      24,721
      2009           .           .           .         640      24,428
      2010           .           .           .         610      24,178
      2011           .           .           .         550      23,852
      2012           .           .           .         459      23,654
      2013           .           .           .         423      23,533
      2014           .           .           .         368      23,442
      2015           .           .           .         333      23,367
      2016           .           .           .         342      23,207
      Sum            .           .           .      12,762     501,105
T2    1996           .           .           .       5,076      65,085
      1997           .           .           .       5,219      70,789
      1998           .           .           .       5,304      76,724
      1999           .           .           .       5,714      83,444
      2000           .           .           .       5,793      89,906
      2001           .           .           .       5,907      96,228
```

	2002	.	.	.	6,198	103,865
	2003	.	.	.	6,530	113,385
	2004	.	.	.	6,440	123,774
	2005	.	.	.	6,728	133,264
	2006	.	.	.	6,838	140,114
	2007	.	.	.	7,066	147,586
	2008	.	.	.	7,008	156,940
	2009	.	.	.	7,688	167,119
	2010	.	.	.	7,870	178,563
	2011	.	.	.	7,982	194,588
	2012	.	.	.	8,457	212,609
	2013	.	.	.	8,731	223,891
	2014	.	.	.	9,164	231,911
	2015	.	.	.	9,546	239,342
	2016	.	.	.	9,741	246,867
	Sum	.	.	.	149,000	3,095,994
Sum	1996	1,194	11,405	12,599	59,955	5,281,172
	1997	1,173	10,757	11,930	59,331	5,302,381
	1998	1,111	11,824	12,935	57,947	5,320,722
	1999	1,005	12,446	13,451	58,814	5,337,072
	2000	988	12,197	13,185	57,156	5,354,934
	2001	1,001	12,244	13,245	57,789	5,374,784
	2002	988	15,356	16,344	57,883	5,393,256
	2003	931	16,819	17,750	57,033	5,407,521
	2004	897	17,010	17,907	54,886	5,421,420
	2005	896	14,684	15,580	54,268	5,436,666
	2006	936	14,112	15,048	54,142	5,454,300
	2007	948	15,530	16,478	54,781	5,476,480
	2008	913	17,400	18,313	53,313	5,505,353
	2009	925	18,482	19,407	53,913	5,533,249
	2010	896	20,571	21,467	53,138	5,556,917
	2011	854	28,743	29,597	51,392	5,580,181
	2012	801	22,799	23,600	51,015	5,601,928
	2013	822	18,186	19,008	50,901	5,625,699
	2014	806	17,241	18,047	50,008	5,657,925
	2015	851	17,625	18,476	51,393	5,700,692
	2016	776	18,521	19,297	51,273	5,730,509
	Sum	19,712	343,952	363,664	1,150,331	115,053,163

We can devise the incidence and mortality rates — from which we see that there is actually quite a difference between the overall mortality rates and the non-DM mortality rates; mainly because of the different age-distributions:

```
> fCtable( tt[,,1:4]/tt[,,c(5,5,5,5)]*10^5, row.vars=1:2, w=9, d=1 )
```

		T1	T2	DM	D
state	P				
noDM	1996	23.0	219.6	242.6	1,044.8
	1997	22.5	206.5	229.0	1,026.1
	1998	21.3	226.5	247.8	995.2
	1999	19.2	238.0	257.2	1,001.8
	2000	18.9	232.7	251.6	966.2
	2001	19.1	233.0	252.1	973.9
	2002	18.8	291.7	310.4	967.4
	2003	17.7	319.2	336.8	943.9
	2004	17.0	322.6	339.6	905.6
	2005	17.0	278.2	295.1	887.1
	2006	17.7	266.8	284.5	881.3

```
      2007      17.9     292.8     310.7     886.5
      2008      17.1     326.8     344.0     857.8
      2009      17.3     346.0     363.3     853.4
      2010      16.7     384.2     400.9     834.1
      2011      15.9     536.1     552.0     799.4
      2012      14.9     424.9     439.8     784.6
      2013      15.3     338.1     353.4     776.2
      2014      14.9     319.1     334.0     749.2
      2015      15.6     324.1     339.8     763.4
      2016      14.2     339.2     353.4     754.3
T1    1996         .         .         .   2,772.3
      1997         .         .         .   2,920.5
      1998         .         .         .   2,963.9
      1999         .         .         .   3,009.6
      2000         .         .         .   3,037.2
      2001         .         .         .   2,954.1
      2002         .         .         .   3,097.3
      2003         .         .         .   3,116.8
      2004         .         .         .   2,830.9
      2005         .         .         .   2,884.9
      2006         .         .         .   2,770.2
      2007         .         .         .   2,791.4
      2008         .         .         .   2,584.9
      2009         .         .         .   2,619.9
      2010         .         .         .   2,523.0
      2011         .         .         .   2,305.9
      2012         .         .         .   1,940.5
      2013         .         .         .   1,797.5
      2014         .         .         .   1,569.8
      2015         .         .         .   1,425.1
      2016         .         .         .   1,473.7
T2    1996         .         .         .   7,799.0
      1997         .         .         .   7,372.6
      1998         .         .         .   6,913.1
      1999         .         .         .   6,847.7
      2000         .         .         .   6,443.4
      2001         .         .         .   6,138.5
      2002         .         .         .   5,967.4
      2003         .         .         .   5,759.1
      2004         .         .         .   5,203.0
      2005         .         .         .   5,048.6
      2006         .         .         .   4,880.3
      2007         .         .         .   4,787.7
      2008         .         .         .   4,465.4
      2009         .         .         .   4,600.3
      2010         .         .         .   4,407.4
      2011         .         .         .   4,102.0
      2012         .         .         .   3,977.7
      2013         .         .         .   3,899.7
      2014         .         .         .   3,951.5
      2015         .         .         .   3,988.4
      2016         .         .         .   3,945.8
```

For the sake of overview in the publications we devise the numbers for the seven 3-year periods too, and this time nicely formatted:

```
> tt <- xtabs( cbind(DM=T1+T2,D,Y=Y*1000) ~ st + sex + gP,
+              data = transform( rt,
```

```
+                                st = Relevel( state, list(DM=2:3), first=FALSE ),
+                                gP = factor( floor((P-1996)/3)*3+1996,
+                                             labels=paste(yf<-1996+0:6*3,
+                                                          yf+2,sep="--") ) ) )
> str(tt)
 'xtabs' num [1:2, 1:2, 1:7, 1:3] 20502 0 16962 0 21901 ...
 - attr(*, "dimnames")=List of 4
  ..$ st : chr [1:2] "noDM" "DM"
  ..$ sex: chr [1:2] "M" "W"
  ..$ gP : chr [1:7] "1996--1998" "1999--2001" "2002--2004" "2005--2007" ...
  ..$    : chr [1:3] "DM" "D" "Y"
 - attr(*, "call")= language xtabs(formula = cbind(DM = T1 + T2, D, Y = Y * 1000) ~ st + sex +
> fCtable( addmargins( tt, 2:3, FUN=list("M+W"=sum,"1996--2016"=sum) ),
+          row.vars=2:3, w=11 )
Margins computed over dimensions
in the following order:
1: sex
2: gP
                st        noDM                                  DM
                            DM           D           Y          DM           D           Y
sex gP
M   1996--1998          20,502      78,885   7,715,200           .       9,076     145,141
    1999--2001          21,901      74,519   7,766,999           .      10,279     179,389
    2002--2004          28,083      71,680   7,810,245           .      11,505     219,973
    2005--2007          26,719      67,787   7,842,954           .      12,144     266,133
    2008--2010          34,118      65,825   7,916,764           .      13,420     314,841
    2011--2013          40,043      61,410   7,956,646           .      15,031     386,881
    2014--2016          31,937      60,230   8,066,473           .      16,613     435,714
    1996--2016         203,303     480,336  55,075,282           .      88,068   1,948,073
W   1996--1998          16,962      80,783   7,908,376           .       8,489     135,558
    1999--2001          17,980      79,691   7,959,048           .       9,270     161,353
    2002--2004          23,918      76,751   7,997,847           .       9,866     194,132
    2005--2007          20,387      72,678   8,029,151           .      10,582     229,208
    2008--2010          25,069      70,084   8,102,806           .      11,035     261,108
    2011--2013          32,162      65,296   8,149,037           .      11,571     315,245
    2014--2016          23,883      62,950   8,234,517           .      12,881     352,422
    1996--2016         160,361     508,233  56,380,782           .      73,694   1,649,027
M+W 1996--1998          37,464     159,668  15,623,576           .      17,565     280,700
    1999--2001          39,881     154,210  15,726,047           .      19,549     340,742
    2002--2004          52,001     148,431  15,808,092           .      21,371     414,105
    2005--2007          47,106     140,465  15,872,105           .      22,726     495,342
    2008--2010          59,187     135,909  16,019,570           .      24,455     575,949
    2011--2013          72,205     126,706  16,105,683           .      26,602     702,126
    2014--2016          55,820     123,180  16,300,990           .      29,494     788,137
    1996--2016         363,664     988,569 111,456,064           .     161,762   3,597,100
> # fCtable( tt[,,1:4]/tt[,,c(5,5,5,5)]*10^5, row.vars=1:2, w=9, d=1 )
```

9.1.1 Arrays of predictions

We set up arrays to hold the predicted incidence and mortality rates from the different models, separately for the two sexes; we are using points for midpoints of age and calendar time categories to identify rates and for boundaries of calendar time categories to identify prevalences:

```
> ht <- function(x) round(c(head(x,3),tail(x,3)),3)
> ht(a.pt)
[1]  0.042  0.125  0.208 99.792 99.875 99.958
> ht(p.pt)
[1] 1996.042 1996.125 1996.208 2039.792 2039.875 2039.958
> ht(t.pt)
[1] 1996.000 1996.083 1996.167 2039.833 2039.917 2040.000
> int ; mean( diff(a.pt) ) ; mean( diff(p.pt) )
[1] 0.08333333
[1] 0.08333333
[1] 0.08333333
```

Here are the arrays to hold the predicted rates.

```
> Lambda <- Mu.nD <- Mu.DM <-
+ NArray( list( a = a.pt,
+                p = p.pt,
+              sex = c("M","W"),
+              mod = c("ap","apc","gam","LCa",
+                      "att","fix","p20","p40","p60") ) )
> str( Lambda )
 logi [1:1200, 1:528, 1:2, 1:9] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 4
  ..$ a  : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" ...
  ..$ p  : chr [1:528] "1996.04166666667" "1996.125" "1996.20833333333" "1996.29166666667" ...
  ..$ sex: chr [1:2] "M" "W"
  ..$ mod: chr [1:9] "ap" "apc" "gam" "LCa" ...
> length( Lambda )
[1] 11404800
```

Note that on dimension 4, we have levels to hold the naive prediction using the natural splines for the AP, APC (and APC with `gam`) and Lee-Carter models, as well for the model with trend attenuation (see below for explanation), `att`. Further we have levels for the models with stronger attenuation and an added time-trend, `atx`, and finally for the simpler scenarios where the rates at the end of 2016 are predicted to increase 2, 2.5 and 3.0% per year, respectively.

9.1.2 Datasets for rate modeling

First we construct simple datasets of identical structure for APC analysis, separately for the three types of transitions we are going to consider. Note that we are resacling the `Y` from person-millenia to person-years. This is all mainly a vehicle for simpler code.

```
> incdat <- transform( subset( rt, state=="noDM" ),
+                      D = T1+T2,
+                      Y = Y * 1000,
+                      A = Ax,
+                      P = Px,
+                      C = Cx )[,c("sex","A","P","C","D","Y")]
> mnDdat <- transform( subset( rt, state=="noDM" ),
+                      Y = Y * 1000,
+                      A = Ax,
```

```
+                       P = Px,
+                       C = Cx )[,c("sex","A","P","C","D","Y")]
> mDMdat <- transform( subset( rt, state!="noDM" ),
+                       Y = Y * 1000,
+                       A = Ax,
+                       P = Px,
+                       C = Cx )[,c("sex","A","P","C","D","Y")]
> str( incdat ) ; str( mnDdat ) ; str( mDMdat )
'data.frame':        8400 obs. of  6 variables:
 $ sex: Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ A  : num  0.667 0.333 0.667 0.333 0.667 ...
 $ P  : num  1996 1997 1997 1998 1998 ...
 $ C  : num  1996 1996 1997 1997 1998 ...
 $ D  : num  0 0 1 0 0 0 1 1 4 0 ...
 $ Y  : num  17894 17576 17340 17746 17089 ...
'data.frame':        8400 obs. of  6 variables:
 $ sex: Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ A  : num  0.667 0.333 0.667 0.333 0.667 ...
 $ P  : num  1996 1997 1997 1998 1998 ...
 $ C  : num  1996 1996 1997 1997 1998 ...
 $ D  : num  28 0 18 1 18 2 11 0 21 0 ...
 $ Y  : num  17894 17576 17340 17746 17089 ...
'data.frame':        345950 obs. of  6 variables:
 $ sex: Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ A  : num  0.667 0.667 0.333 0.667 0.667 ...
 $ P  : num  1997 1999 2000 2000 2000 ...
 $ C  : num  1997 1999 1999 2000 2000 ...
 $ D  : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y  : num  0.1431 0.0226 0.0767 0.3793 0.1758 ...
> save( incdat, mnDdat, mDMdat, file="../nydata/rate-dat.Rda" )
```

9.2 Models for incidence and mortality rates

There are two immediate alternatives to the age-period-cohort models on the log-scale; age-period-cohort models on the additive scale and Lee-Carter models. These may provide a better fit to data (some of the expanded Lee-Carter models necessarily do). Further, we may fit the age-period-cohort model using `gam` from the `mgcv` package.

For now, however we proceed with APC-models for the incidence and mortality rates, but also using the `gam` facility.

9.3 Incidence rates

We now derive the location of the knots for the age, period and cohort terms in the model.

```
> nk.p <- 9
> c(nk.a,nk.p,nk.c)
[1] 8 9 8
> ( ki.a <- with( incdat, quantile( rep(  A,D), qn(nk.a) ) ) )
   6.25%   18.75%   31.25%   43.75%   56.25%   68.75%   81.25%   93.75%
36.33333 48.66667 54.66667 60.33333 64.66667 69.33333 74.66667 82.66667
```

```
> ( ki.p <- with( incdat, quantile( rep(P  ,D), qn(nk.p) ) ) )
5.555556% 16.66667% 27.77778% 38.88889%       50% 61.11111% 72.22222% 83.33333% 94.44444%
 1997.667  2000.667  2003.333  2005.667  2008.333  2010.333  2011.667  2013.667  2015.667
> ( ki.c <- with( incdat, quantile( rep(P-A,D), qn(nk.c) ) ) )
   6.25%   18.75%   31.25%   43.75%   56.25%   68.75%   81.25%   93.75%
1922.333 1931.667 1938.333 1943.333 1947.667 1953.333 1960.333 1972.667
```

The model we set up is an age-period-cohort model with these three terms in it. As we are only going to use the model for predictions we need not bother about parametrization issues, so it is not an issue that the model we fit is formally over-parametrized. However we want to extract the average trend from the APC-model, so we also fit the model with the parametrization that allows us to extract the trend.

```
> m.inc.ap <- glm( D ~ Ns(A,kn=ki.a) + Ns(P,kn=ki.p),
+                      offset = log( Y ),
+                      family = poisson,
+                        data = subset(incdat,sex=="M") )
> m.inc.aPC <- update( m.inc.ap, . ~ . - Ns(P  ,kn=ki.p) + I(P) +
+                                 detrend( Ns(P  ,kn=ki.p), P  , Y ) +
+                                 detrend( Ns(P-A,kn=ki.c), P-A, Y ) )
> m.inc.apc <- update( m.inc.ap, . ~ . + Ns(P-A,kn=ki.c) )
> m.inc.gam <- gam( D ~ s(A) + s(P) + s(C),
+                       offset = log( Y ),
+                       family = poisson,
+                         data = subset(incdat,sex=="M") )
> c( m.inc.apc$deviance, m.inc.aPC$deviance )
[1] 7077.282 7077.282
> f.inc.ap  <- update( m.inc.ap , data = subset(incdat,sex=="W") )
> f.inc.apc <- update( m.inc.apc, data = subset(incdat,sex=="W") )
> f.inc.gam <- update( m.inc.gam, data = subset(incdat,sex=="W") )
> f.inc.aPC <- update( m.inc.aPC, data = subset(incdat,sex=="W") )
```

The average annual trends in incidence from the multiplicative models:

```
> inc.chg <- rbind( ci.exp(m.inc.aPC,subset="I\\(P")-1,
+                   ci.exp(f.inc.aPC,subset="I\\(P")-1 )*100
> rownames( inc.chg ) <- c("DM incidence change     Men",
+                          "                       Women")
> round( inc.chg, 2 )
                           exp(Est.) 2.5% 97.5%
DM incidence change     Men      2.95 2.82  3.09
                      Women      2.79 2.64  2.93
```

The average increase is similar in women and men, but the period effect is massively non-linear, so these summary figures are not really informative, see the comparative figure with the mortality rates below.

9.3.1 Early and late trends in incidence rates

In order to get a fuller picture of the rate trens we replace the linear effect of `P` with a linear spline with a single knot at 2011:

```
> m.inc.aPls <- update( m.inc.aPC, . ~ Ns(A,kn=ki.a) + pmin(P-2011,0) + pmax(P-2011,0) )
> f.inc.aPls <- update( m.inc.aPls, data = subset(incdat,sex=="W") )
> inc.chg <- rbind( ci.exp(m.inc.aPC, subset="I\\(P")-1,
+                   ci.exp(m.inc.aPls,subset="pm"   )-1,
+                   ci.exp(f.inc.aPC, subset="I\\(P")-1,
+                   ci.exp(f.inc.aPls,subset="pm"   )-1)[]*100
> rownames( inc.chg ) <- c("DM incidence change Men total:",
+                          "                      pre-2011",
+                          "                     post-2011",
+                          "                   Women total:",
+                          "                      pre-2011",
+                          "                     post-2011")
> round( inc.chg, 2 )
                                exp(Est.)  2.5% 97.5%
DM incidence change Men total:      2.95  2.82  3.09
                     pre-2011       4.18  4.07  4.30
                    post-2011      -4.45 -4.75 -4.16
                  Women total:      2.79  2.64  2.93
                     pre-2011       3.79  3.67  3.92
                    post-2011      -4.67 -5.00 -4.34
```

Thus we see that there is substantial change in rates around 2011. The change after 2011 is based on quite a short period of time, so presumably quite influenced by the precise location of the knot.

9.3.2 Incidence rate predictions

We want the predicted incidence rates at a grid of points suitable for the calculations of predicted prevalences — essentially for the construction of transition probabilities. We make the predictions for all combinations of `a.pt` and `p.pt`.

All the predictions should be in units of the interval length chosen for calculations, or more precisely we predict cumulative rates over small intervals. Thus we use a prediction data frame with the person-years-variables set to `int`, the length of the step interval we shall use for projections.

Note that the `Lambda` array was set up with age before period, so that the column-major storage of arrays conforms with the predictions obtained using the following `nd` as `newdata`. Note that we are computing the rates at the boundaries of the age-classes (hence the term `+int/2`) but at the midpoints of the periods:

```
> ndn <- function(obj,dno) as.numeric(dimnames(obj)[[dno]])
> nd <- transform( data.frame( expand.grid( A = ndn(Lambda,1),
+                                           P = ndn(Lambda,2) ),
+                              Y = int ),
+                  C = P - A )
> str( nd )
'data.frame':        633600 obs. of  4 variables:
 $ A: num  0.0417 0.125 0.2083 0.2917 0.375 ...
 $ P: num  1996 1996 1996 1996 1996 ...
 $ Y: num  0.0833 0.0833 0.0833 0.0833 0.0833 ...
 $ C: num  1996 1996 1996 1996 1996 ...
> head( nd )
           A        P          Y        C
1 0.04166667 1996.042 0.08333333 1996.000
```

```
2 0.12500000 1996.042 0.08333333 1995.917
3 0.20833333 1996.042 0.08333333 1995.833
4 0.29166667 1996.042 0.08333333 1995.750
5 0.37500000 1996.042 0.08333333 1995.667
6 0.45833333 1996.042 0.08333333 1995.583
> fCp( c( prod( dim(Lambda)[1:2] ), length(Lambda),
+         nrow(nd), nrow(nd)*prod( dim(Lambda)[-(1:2)] ) ) )
[1]   633,600  11,404,800   633,600  11,404,800
> Lambda[,,"M","ap" ] <- ci.pred( m.inc.ap , newdata=nd )[,1]
> Lambda[,,"W","ap" ] <- ci.pred( f.inc.ap , newdata=nd )[,1]
> Lambda[,,"M","apc"] <- ci.pred( m.inc.apc, newdata=nd )[,1]
> Lambda[,,"W","apc"] <- ci.pred( f.inc.apc, newdata=nd )[,1]
> Lambda[,,"M","gam"] <- ci.pred( m.inc.gam, newdata=nd )[,1]*int
> Lambda[,,"W","gam"] <- ci.pred( f.inc.gam, newdata=nd )[,1]*int
> # Lambda[,,"M","LCa"] <- predict.LCa( m.inc.LCa, newdata=nd )[,1]
> # Lambda[,,"W","LCa"] <- predict.LCa( f.inc.LCa, newdata=nd )[,1]
```

Note that we multiply the `gam` predictions by `inc` because the prediction machinery for `gam` ignores the offset when it is given as an argument instead of a model term. This may be a facility.

Thus we have the incidence rates that we need for two sexes, and for two different modeling approaches. Note that we have computed the rates in units of `int`=0.08333. This renders the numbera as a close approximation to the cumulative rate over the interval, and since this is small also a close approximation to the cumulative risk over the interval. Which is what we shall need for the predictions.

9.3.3 Calendar time trends

We have modeled the diabetes incidence rates over the Lexis diagram (1996–2017)×(0–100) with three different models. We can show how the rates in ages 20 through 90 evolve over time under each of the models:

```
> pl.pa <-
+ function( mod, xl=c(1996,2017) )
+    {
+ for( sx in dimnames(Lambda)[["sex"]] )
+    {
+    wh.a <- 2:9*120
+    matplot( p.pt, t(Lambda[wh.a,,sx,mod])/int*1000,
+             type="l", lwd=2, lty=1,
+             col=if(sx=="M") "blue" else "red",
+             xlab="", ylab="", xlim=xl,
+             ylim= c(0.1,20), yaxs="i", xaxt="n", yaxt="n", log="y" )
+    axis( side=1, at=seq(1995,2040,5 ) )
+    axis( side=1, at=1996:2040, tcl=-0.3, labels=NA )
+    if( sx=="M" )
+      {
+      axis( side=2)
+      axis( side=2, at=1:19, labels=NA, tcl=-0.3 )
+      }
+   text( 2017, Lambda[wh.a,21*12,sx,mod]/int*1000, 2:9*10, adj=1, font=2 )
+    }
+ mtext( "Date of follow-up", 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 )
+    }
> par( mfrow=c(3,2), mar=c(3,0,0,0), oma=c(0,4,1,0),
+      mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pl.pa("ap")
> pl.pa("apc")
> pl.pa("gam")
> mtext( "DM incidence rates per 1000 PY",
+        side=2, outer=TRUE, line=3, las=0, cex=0.67 )


> par( mfrow=c(3,2), mar=c(3,0,0,0), oma=c(0,4,1,0),
+      mgp=c(3,1,0)/1.6, las=1, bty="n" )
> pl.pa("ap" , xl=c(1996,2040) )
> pl.pa("apc", xl=c(1996,2040) )
> pl.pa("gam", xl=c(1996,2040) )
> mtext( "DM incidence rates per 1000 PY",
+        side=2, outer=TRUE, line=3, las=0, cex=0.67 )
```

Now, there is a clear trend in the rates at the end; we can show the trend in different ages by extracting the last few dates from the `Lambda` array:

```
> ( ell <- dimnames(Lambda)[[2]][c(240,253)] )
[1] "2015.95833333333" "2017.04166666667"

> ( ell <- diff( as.numeric( ell ) ) )
[1] 1.083333

> tr.ap  <- (exp(log(Lambda[,253,,"ap" ]/Lambda[,240,,"ap" ])/ell)-1)*100
> tr.apc <- (exp(log(Lambda[,253,,"apc"]/Lambda[,240,,"apc"])/ell)-1)*100
> tr.gam <- (exp(log(Lambda[,253,,"gam"]/Lambda[,240,,"gam"])/ell)-1)*100
> matplot( as.numeric(dimnames(Lambda)[[1]]), cbind( tr.ap, tr.apc, tr.gam ),
+          las=1, type="l", lty=rep(c("solid","21","solid"),each=2),
+          lend=1, lwd=c(1,1,3,3,3,3), col=c("blue","red"),
+          ylim=range(c(0,tr.ap,tr.apc,tr.gam)),
+          xlab="Age", ylab="Annual change in rates (%)" )
> abline( h=0, col="gray" )
```

9.4 Attenuated extrapolation

Instead of using the naive extrapolation as above where we blindly just prolong the linear trajectories from the natural splines (for period and cohort), we may dampen the trend derived from the naive application of the natural splines. The attenuation factor (explained below) is traditionally (*i.e.* in cancer epidemiology [3, 4]) set to 0.92 per year (corresponding to a halving of the drift approximately every 8.3 years).

In the subsequent implementation we shall parametrize the damping factor by the half-time of the slope, thus instead of d use h where: $d^h = 0.5 \Leftrightarrow d = 0.5^{1/h}$.

In order to do this we will need to *have* a drift to halve. To this end we may choose a parametrization of the form:

$$\log(\lambda(a,p)) = f(a) + g(p) + h(p-a) + \beta p$$

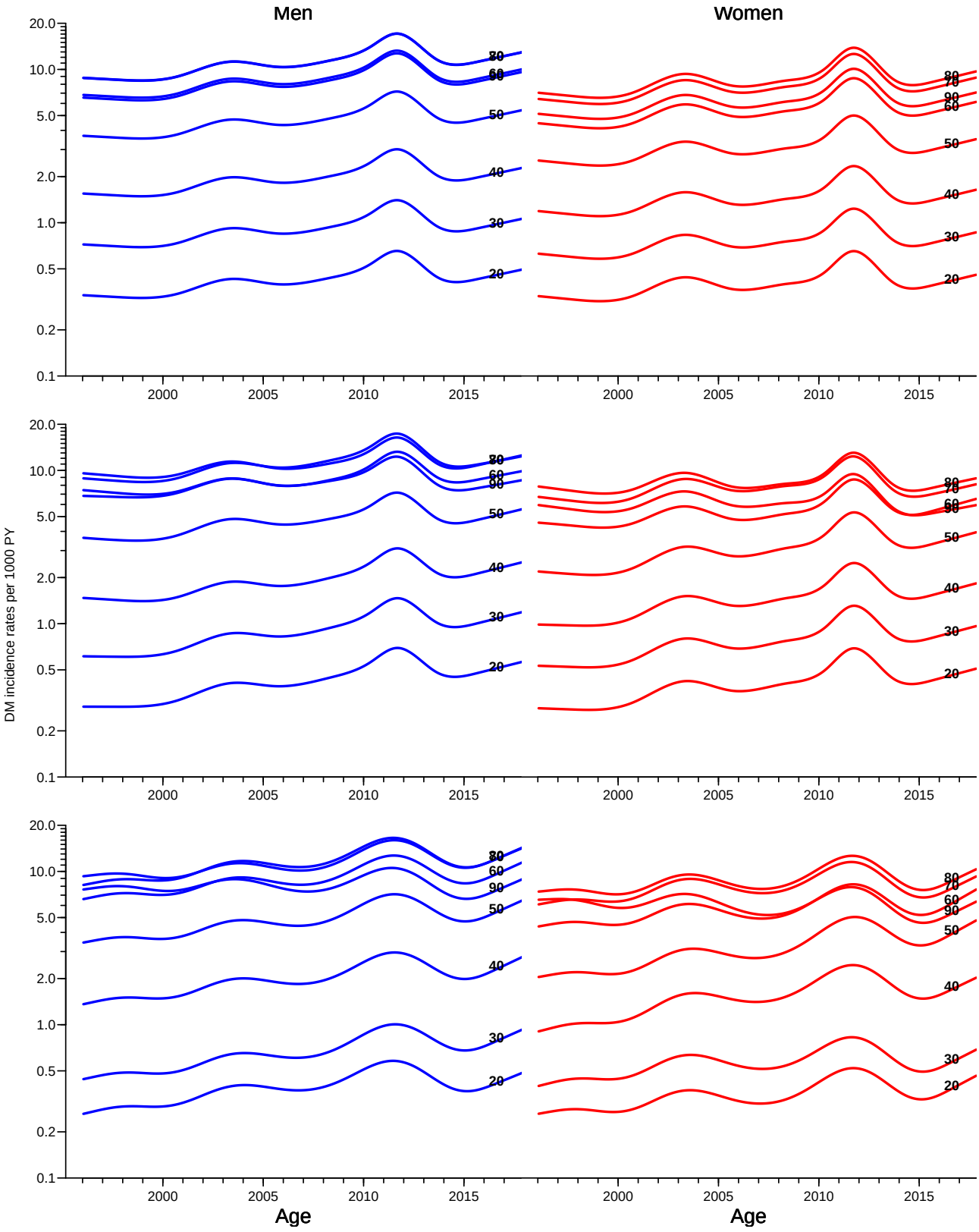


Figure 9.1: *Timetrend in diabetes incidence rates in Denmark for different ages. Models from top are AP, APC and gam. Blue: man, red: women. Shaded areas are 95% confidence intervals.* `./graph/pr-rates-per-age`

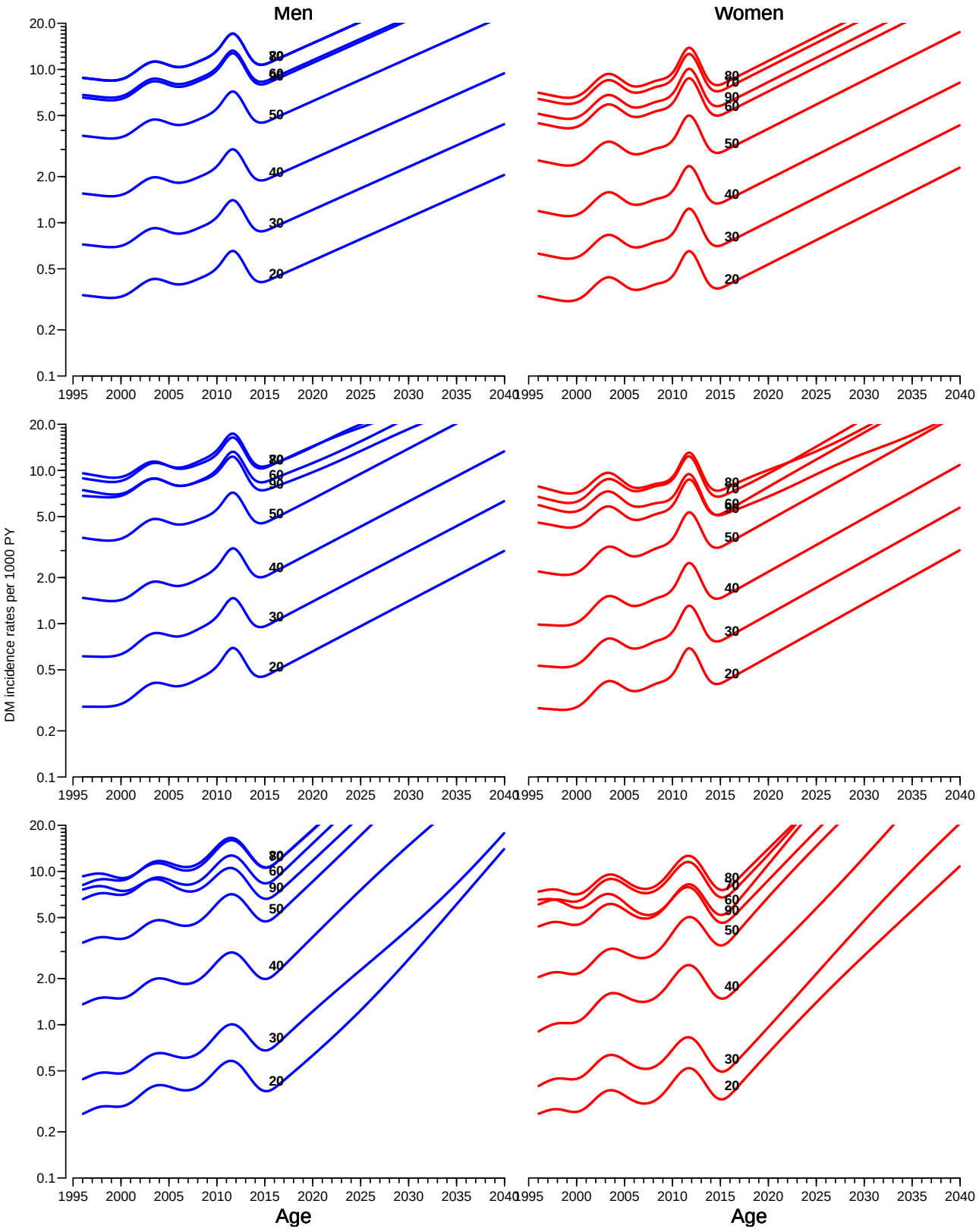


Figure 9.2: *Timetrend in diabetes incidence rates in Denmark for different ages with naive predictions to 2040. Models from top are AP, APC and gam. Blue: man, red: women.* `./graph/pr-rates-per-age-full`

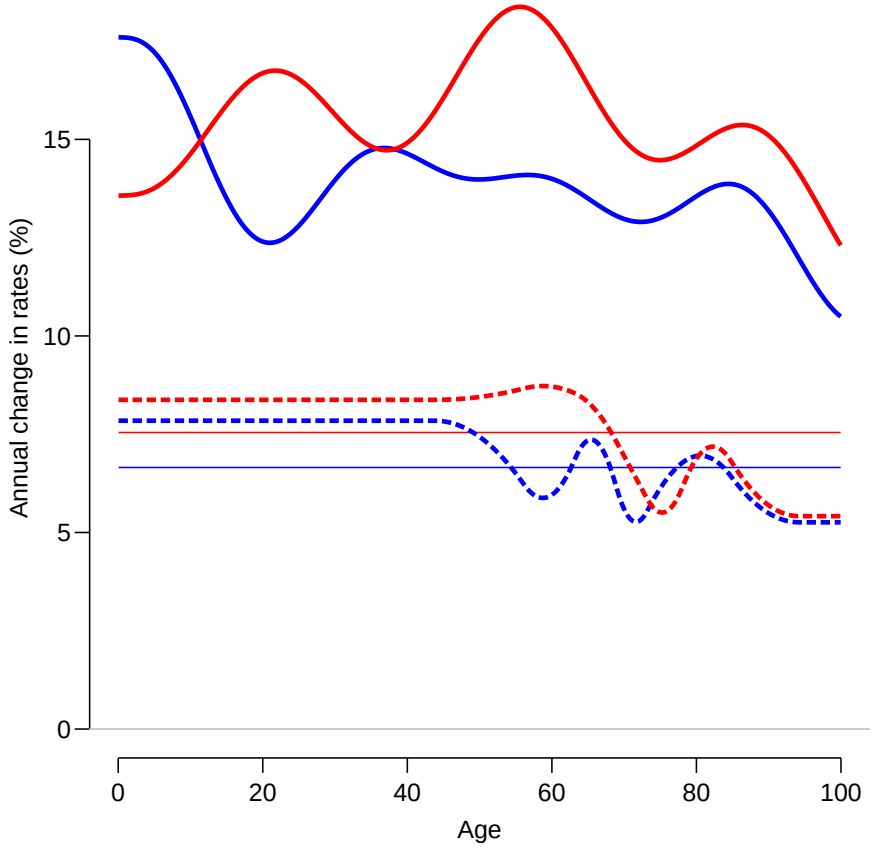


Figure 9.3: *Estimated trend in diabetes incidence rates by age over the year 2016, red curves are women, blue men, the dotted curves are from the APC-model, the thick full lines from the* `gam` *model, and the thin horizontal lines from the AP model.* `./graph/pr-rates-last-trend`

where g and h are chosen to be flat on average. This is generally in the literature thought to be a unique operation, however it depends on the choice of an inner product in the space of the linear predictor [1][1].

Now suppose we have a parametrization using natural splines [4] with extracted drift for some choice of inner product (that is the model matrix columns defining $g(p)$ are orthogonal to the intercept and p, and similarly for $h(p-a)$[1]). A natural extrapolation beyond the last observation date p_x, say would be to use for $p > p_x$:

$$\log(\lambda(a,p)) = f(a) + g(p_x) + h(p_x - a) + \beta p_x + \beta d^{p-p_x}(p - p_x)$$

where d is a damping factor, such that the drift at date p after p_x is βd^{p-p_x}.

9.4.1 Theory

The following is an explanation of the potential problem with the approach of Sasieni (as I read it), but it is only an explanation of why a different approach is needed in order to obtain a damping of future drifts from APC-models that are invariant under reparametrizations of the APC-model.

Exponential damping of a drift

Arithmetically, the attenuation or "damping" machinery works as follows in *continuous* time:

If we model age, period and cohort terms by natural splines, *i.e.* using splines that are linear beyond the outer knots, then these easily lend themselves to prediction beyond data, simply by extending the linear part of the natural spline, see *e.g.* Rutherford *et al.*[4].

Now suppose the slope of the period or cohort effect at the end is β, but that it would be an exaggeration to continue the period effect indefinitely at a slope of β (a prediction of the form $f(t) = k + \beta t$). Therefore we choose a *damping* factor, d, say, such that the *slope* of the effect at time t is not β, but rather βd^t. If $d < 1$ it means that the slope will gradually approach 0 as $t \to \infty$, the faster the smaller d is, so as mentioned above $d = 0.92$ implies that $d^{8.3} = 0.5$ and $d = 0.88$ implies that $d^{5.4} = 0.5$. Subsequently we shall parametrize by the halving time instead of d.

For the damped effect itself, $f(t)$, we have:

$$f'(t) = \beta d^t \quad \Leftrightarrow \quad f(t) = k + \beta d^t / \log(d) \quad \Rightarrow \quad f(0) = k + \beta / \log(d)$$

Solving for k, we get:

$$k = f(0) - \beta / \log(d) \quad \Rightarrow \quad f(t) = f(0) - \frac{\beta}{\log(d)} + \frac{\beta d^t}{\log(d)} = f(0) + \beta \frac{d^t - 1}{\log(d)}$$

where $f(0)$ is the value of the period or cohort term at the beginning of the prediction interval.

Note that the term $(d^t - 1)/\log(d)$ is always non-negative for $0 < d < 1$.

[1] Peter Sasieni [5] refer to this model but does not explain how the drift (β) was extracted — the parametrization with a separte drift is not unique, though apparently believed so (?).

Implementation in APC-models

Suppose the fitted APC model is parametrized as:

$$\log(\lambda(a,p)) = f(a) + g(p) + h(p-a) + \beta p$$

where the last term is brought about by some sort of projection [1], effectively leaving $g(p)$ and $h(p-a)$ as "de-trended", that is with overall average slope 0 in some (in this context, unspecified) sense.

If we assume that g and h are parametrized by natural splines, then $g(p)$ is linear beyond the last knot which we shall assume to be smaller than the end of data (well, really the starting point of the prediction period), p_x, say. Now, any reparametrization involving the period / cohort slopes will give a model of the form (for an arbitrary δ):

$$\begin{aligned}\eta_0(a,p) = \log\big(\lambda(a,p)\big) &= \big(f(a) - \delta a\big) + \big(g(p) + \delta p\big) + \big(h(p-a) - \delta c\big) + \beta p && (9.1)\\ &= \big(f(a) - \delta a\big) + g(p) + \big(h(p-a) - \delta c\big) + (\beta+\delta)p && (9.2)\end{aligned}$$

If a prediction of future rates is going to depend non-linearly on the time since the prediction point, for example by specifying the slope to be the extracted slope multiplied by d^{p-p_x} for some damping factor $0 < d < 1$, the prediction will depend in the parametrization chosen in the following sense:

If we use the parametrization 9.1 to construct predicted rates, the damped version of the log-rates would be (using the notation $(x)_+ = \max(x,0)$):

$$\eta_1(a,p) = \eta_0(a,p) - \beta(p-p_x)_+ + \frac{\beta(d^{(p-p_x)_+} - 1)}{\log(d)}$$

whereas if we use the parametrization 9.2 they would be:

$$\begin{aligned}\eta_2(a,p) &= \eta_0(a,p) - (\beta+\delta)(p-p_x)_+ + (\beta+\delta)\frac{d^{(p-p_x)_+} - 1}{\log(d)}\\ &= \eta_1(a,p) - \delta(p-p_x)_+ + \delta\frac{d^{(p-p_x)_+} - 1}{\log(d)}\end{aligned}$$

So the prediction with a damping factor (or for that matter any other replacement of βp with a non-linear function of p) will not be invariant under reparametrization of the original model.

Resorting to the postulated uniqueness of the drift as claimed by e.g. Holford [2], will not solve the problem, it will just choose a particular way of extracting the drift [1]. Other ways of extracting the drift will give different results. It all boils down to replacing the predictions based on extrapolation of the period and cohort term by the linear parts of the natural splines by adding the term

$$\beta\big(-(p-p_x)_+ + (d^{p-p_x} - 1)/\log(d)\big)$$

for some chosen β.

The good news is that different reasonable ways of extracting the trend do not normally give dramatically different values for the drift β.

The bad news is that the extracted slope may not be anywhere near any slope of rates at the beginning of the prediction interval.

Thus it seems that the reliance on the particular form of parametrization is not a viable path, so here is a more directly data-driven approach, that for a given APC-model is parametrization invariant.

9.4.2 A discrete time shortcut

The following is an empirical approach to adjust rates predicted into the future. We use a damping mechanism, taking an approach that does not rely on any particular mathematical form of the predictions, but merely on the predictions being available in suitably small intervals.

Suppose you have prediction of future rates (or log-rates) $\lambda(a, p)$ from an APC-model (well, this goes for any model) — values for occurrence rates in the period-direction.

A slope-attenuation can be numerically implemented by using the empirical gradients of the predictions, so suppose that for a *fixed* value of age (a) the rates are in the vector `f` and the corresponding dates (p) in the vector `t`. The empirical slopes between successive time points is then simply `diff(f)/diff(t)`, and we could then attenuate these directly by multiplying them by d^t where d is the chosen damping factor and t is the midpoint of the interval. Here is a simple illustration of how it goes:

```
> t <- 0:100/20
> f <- 0.2*t + sin(t) + 0.2*(t)^2
> f <- f/diff(range(f))
```

Then we need the midpoints of the observation intervals:

```
> dd <- 0.8
> dt <- diff(t)
> mt <- t[-1] - dt/2
>  df <- diff(f) / diff(t)
> ddf <- df * dd^mt
> iof <- c( f[1], f[1] + cumsum( df)*dt )
> idf <- c( f[1], f[1] + cumsum(ddf)*dt )
```

Now this is easily implemented in a function which takes the function values `f`, times `t` and damping factor as arguments. What remains to be seen is that this method will produce results that are (at least approximately) invariant under choice of step-size; so we take every 4^{th} and see how it goes:

```
> t <- t[0:25*4+1]
> f <- f[0:25*4+1]
> dt <- diff(t)
> mt <- t[-1] - dt/2
>  df <- diff(f) / diff(t)
> ddf <- df * dd^mt
> iof <- c( f[1], f[1] + cumsum( df)*dt )
> idf <- c( f[1], f[1] + cumsum(ddf)*dt )
```

We can then plot the results from the two approaches together:

```
> # intervals of 0.2
> t <- 0:100/20
> f <- 0.2*t + sin(t) + 0.2*(t)^2
> f <- f/diff(range(f))
> plot( f ~ t, type="l", lwd=8 )
> dd <- 0.8
> dt <- diff(t)
> mt <- t[-1] - dt/2
>  df <- diff(f) / diff(t)
> ddf <- df * dd^mt
```

```
> iof <- c( f[1], f[1] + cumsum( df)*dt )
> idf <- c( f[1], f[1] + cumsum(ddf)*dt )
> matlines( t, cbind( iof, idf),
+           type="l", lwd=4, lty=1, col=c("red","forestgreen") )
> # intervals of 1
> t <- t[0:5*20+1]
> f <- f[0:5*20+1]
> dt <- diff(t)
> mt <- t[-1] - dt/2
>  df <- diff(f) / diff(t)
> ddf <- df * dd^mt
> iof <- c( f[1], f[1] + cumsum( df)*dt )
> idf <- c( f[1], f[1] + cumsum(ddf)*dt )
> matlines( t, cbind( iof, idf),
+           type="l", lwd=4, lty=1, col=gray(0.7) )
```

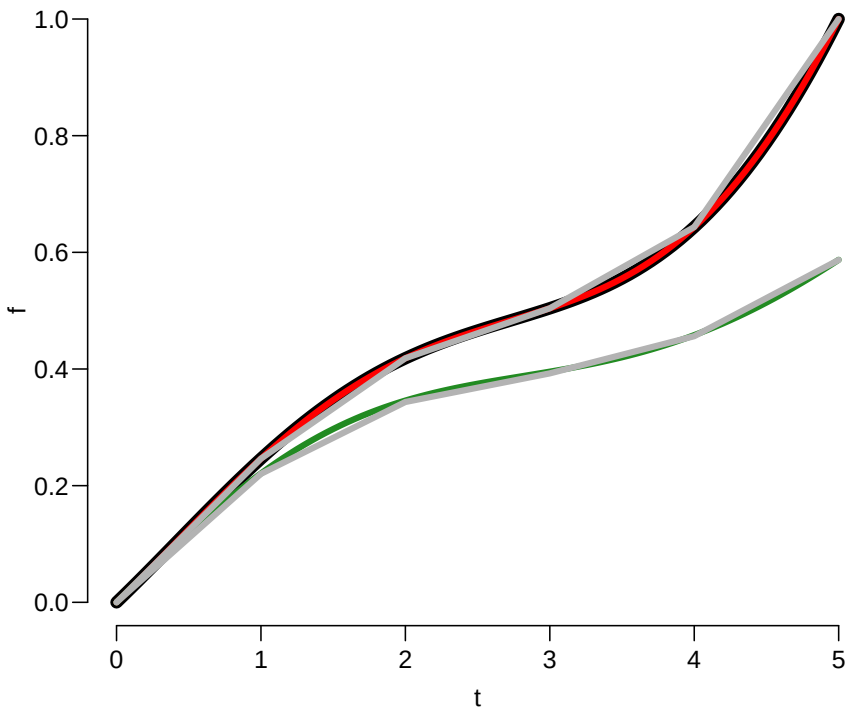


Figure 9.4: *An "arbitrary" curve and the attenuated version, using an attenuation factor of 0.8 (per t unit). The curves are constructed using steps of length 0.2 or 1 (the latter overlaid in gray).* `./graph/pr-rates-att-example`

Indeed the resulting curves coincide at the points where the coarser is defined, although a mathematical calculation would show that they are not exactly the same.

Adding a drift

For the diabetes incidence we have observed that the incidence rates show a dramatically increasing tendency over the last year of observation ($\approx 15 - -20\%$/year), hence we may want not only to investigate a scenario where rates are kept or attenuated to constant, but also one where we simply let the rates increase by a certain amount, say 4% per year. This is only going to be used for the incidence rates as a sensitivity analysis.

To this end we update the damping function briefly outlined above by allowing adding a trend (drift) in calendar time on top of the attenuated prediction; we phase it in quadratically over a period of ℓ, by the function q, a parabola with slope 0 at 0 and slope δ at *ell*, and a linear function with slope δ beyond ℓ, that is:

$$q(t) = \begin{cases} 0 < t < \ell & : \ \big(\delta/(2\ell)\big)t^2 \\ \ell < t < \infty & : \ -\delta\ell/2 + \delta t \end{cases}$$

Thus we see that $q(0) = 0$, and using the first line of the definition: $q(\ell) = \big(\delta/(2\ell)\big)\ell^2 = \delta\ell/2$, which is also obtained using the second line. Moreover, $q'(t) = t\delta/\ell|_{t=\ell} = \delta$.

In R-code this function becomes:

```
> qs <-
+ function( t, ell, delta ) ifelse( t < ell, delta / ell / 2 * t^2,
+                                             delta * t - delta * ell / 2 )
```

which is incorporated in a damping function below.

Implementation

We implement this attenuation and slope addition in a function `damp` which takes 6 arguments:

`f` — a vector of predicted function values (rates or log-rates) to be modified by damping and/or addition of a trend

`t` — an ordered vector of time points where `f` is given. Need not be equidistant. Note that `t-t[1]` is used as exponent to the damping factor, so results are invariant under translation of `t`.

`h` — a scalar, the halving time for the slope. In the function it is converted to a damping factor which will be elevated to the power of `t`, thus dependent on the scaling of `t`: For halving time h we have $d^h = 0.5 \Leftrightarrow d = 0.5^{1/h}$.

`delta` — scalar; the extra slope added to the predictions, beyond `ell` (`t`≥`ell`), before `ell` the addition is a quadratic starting at 0 and a slope fitting with the linear at `ell`. This is an additive factor, so a 10% increase per unit of `t` is obtained by `delta=0.1`.

`ell` — scalar; the run-in interval (on the `t`-scale) for the extra slope.

`logf` — logical indicating whether the supplied `f` represent log-rates or rates. In any case the attenuation is made on the log-rate scale.

With this, a value of 0 for `h` produces an immediately flat (constant) modified curve, corresponding to a fixing of rates at `t` = 0. Likewise a choice of 0 for the interval length `ell` corresponds to an immediate start of an added slope of `delta`:

```
> damp <-
+ function( f, t, h, delta = 0,      # added slope (% per t unit),
+                      ell = 0,      # phase-in interval for added slope
+                     logf = FALSE ) # is f a vector of log-rates
+ {
+ # all operations are on log-rates so if we have rates make them log
```

```
+ if( !logf ) f <- log( f )
+ # compute the damping factor from half-time
+ d <- 0.5^(1/h)
+ # make sure t start at 0
+ t <- t - t[1]
+ # difference between timepoints of prediction
+ dt <- diff(t)
+ # midpoints of intervals
+ mt <- t[-1] - dt/2
+ # slopes in each interval
+ dfdt <- diff(f) / dt
+ # attenuated slopes
+ atdf <- dfdt * d^mt
+ # function values after attenuating the slope
+ idf <- f[1] + cumsum(c(0,atdf))*dt
+ # remember delta is taken as being in % per t
+ delta <- delta/100
+ # add the extra slope to this
+ idf <- idf + ifelse( t < ell, delta/(2*ell)*t^2,
+                              delta*(t-ell/2) )
+ if( !logf ) idf <- exp( idf )
+ idf
+ }
```

We can illustrate the damping effect in different ways, firstly, the time it takes to reduce the slope to say, 50, 10 and 1% (ζ, say) of the original one, is illustrated by simply solving:

$$d^t = \zeta \quad \Leftrightarrow \quad t\log(d) = \log(\zeta) \quad \Leftrightarrow \quad t = \log(\zeta)/\log(d)$$

This is the left panel in figure reffig:damp-ex; the other one illustrates the resulting damped / amended curves relative to an arbitrary constant slope:

```
> par( mfrow=c(1,2), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> clr <- rainbow(3)
> d <- seq(0,1,,200)
> zeta <- c(0.5,0.1,0.01)
> matplot( d, outer( d, zeta, function(d,zeta) log(zeta)/log(d) ),
+          type="l", lwd=4, lty=1, col=clr,
+          ylim=c(0,25), xlab="Damping factor",
+          ylab=paste( "Time to reduction to ",
+                      paste( round(zeta*100,1), collapse=", "),
+                      "%, respectively", sep="" ) )
> abline( v=c(0.92, 0.88, 0.7) )
> abline( h=0:10, lty=2, col=gray(0.8) )
>  axis( at=c(0.92, 0.88, 0.7), las=2, side=1 )
> clr <- c("black",rainbow(7))
> tt <- seq( 0,25,0.1)
> ff <- 12.5 - 0.4 * tt
> t0 <- 8
> t  <- (tt-t0)[tt>=t0]
> f  <-      ff[tt>=t0]
> plot( tt, ff, lty=1, lwd=5, type="l", ylim=c(2,12),
+               xlab="Time", ylab="Damped effect")
> matlines( t+t0, cbind( f, damp(f,t,h=5),
+                           damp(f,t,h=Inf),
+                           damp(f,t,h=10,delta=5,ell=5),
+                           damp(f,t,h=2  ,delta=5,ell=5),
```

```
+                           damp(f,t,h=2),
+                           damp(f,t,h=2  ,delta=-5,ell=5) ),
+          lty=1, lwd=c(5,rep(3,6)), type="l", col=clr,
+          xlab="Time", ylab="Damped effect")
> text( 5, 5.5-0:6/2, c( "Half-time",
+                        formatC( c(5,Inf,10,2,2,2), format="f", digits=2 ) ),
+                     font=2, col=clr, adj=1 )
> text( 7, 5.5-0:6/2, c( "Added slope / yr",
+                        formatC( c(0,0,1/20,abs(1:-1)/20), format="f", digits=2 ) ),
+                     font=2, col=clr, adj=0 )
> text( 6.9, 5.5-6/2, "-", font=2, col=clr[7], adj=1 )
> segments( c(t0,t0+5), 6,
+           c(t0,t0+5), 12 )
```

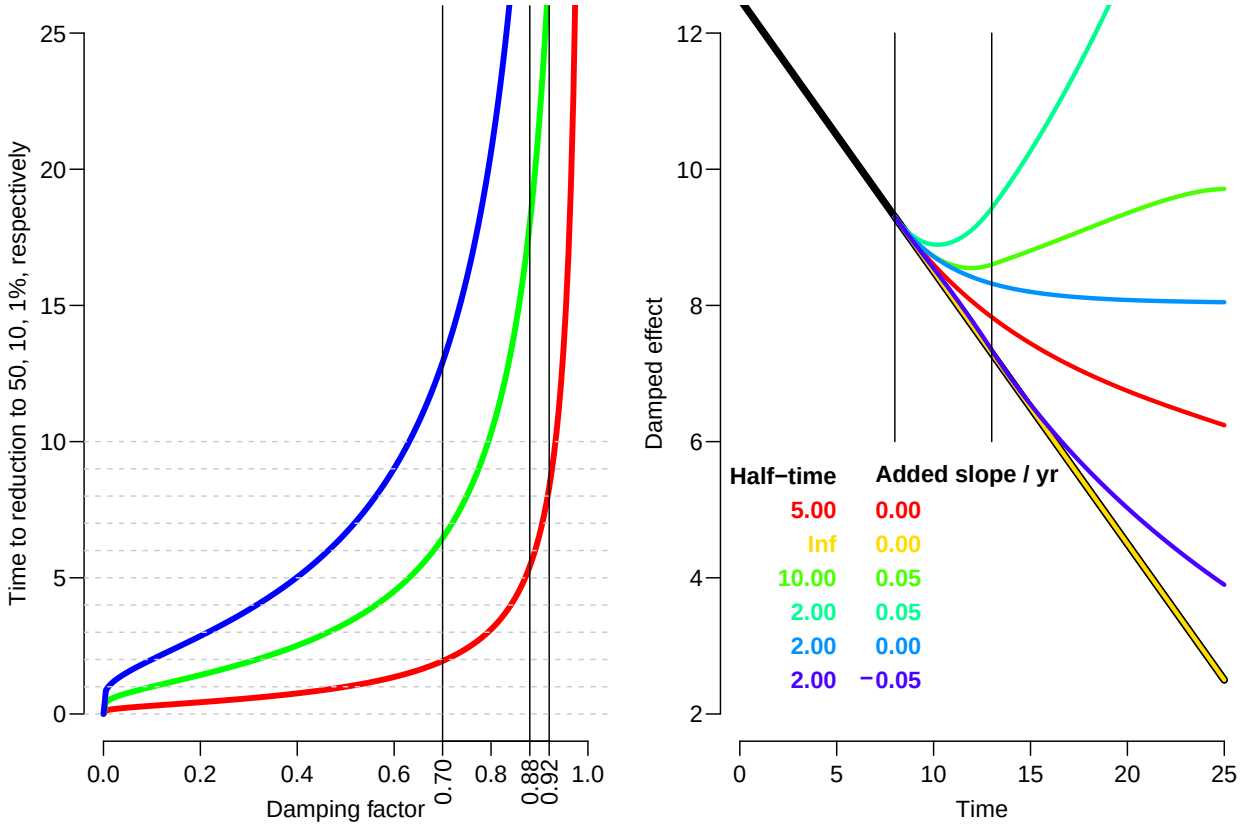


Figure 9.5: *Illustration of the* `damp` *function for attenuation of effects and addition of linear terms for various combinations of the two. The two vertical black lines indicate the starting point of the attenuation and the end of the phase-in of the added slope.*./graph/pr-rates-damp-ex

So for the incidence array `Lambda` we would do something like:

```
> str( Lambda )
 num [1:1200, 1:528, 1:2, 1:9] 6.14e-06 6.18e-06 6.22e-06 6.26e-06 6.30e-06 ...
 - attr(*, "dimnames")=List of 4
  ..$ a  : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" ...
  ..$ p  : chr [1:528] "1996.04166666667" "1996.125" "1996.20833333333" "1996.29166666667" ...
  ..$ sex: chr [1:2] "M" "W"
  ..$ mod: chr [1:9] "ap" "apc" "gam" "LCa" ...
```

```
> # where is the future
> wh.p <- 253:528
> fp.pt <- p.pt[wh.p]
> fp.pt[1:3]
[1] 2017.042 2017.125 2017.208
> rev( rev(fp.pt)[1:3] )
[1] 2039.792 2039.875 2039.958
> for( ia in dimnames(Lambda)[['a']] )
+ for( is in dimnames(Lambda)[['sex']] )
+    {
+    # Compute the damped values along the period dimension
+    # wh.p is the only
+          Lambda[ia,wh.p,is,"att"] <-
+ damp( f = Lambda[ia,wh.p,is,"apc"],
+       t = fp.pt,
+       h = 2,
+   delta = 0,
+     ell = 0 )
+     }
```

The outer loop over age could be replaced by a simple `apply`; further we will also make predictions with added slopes of 2, 4 and 6% per year, but first we need models for the mortality rates.

9.5 Mortality rates

9.5.1 Diabetes patients

First we fix the position of knots by age, period and cohort as we did for the incidence rates, and then we fit the same set of models, and make the same set of predictions, and put in a similarly defined array:

```
> ( kmd.a <- with( mDMdat, quantile( rep(A  ,D), qn(nk.a) ) ) )
   6.25%   18.75%   31.25%   43.75%   56.25%   68.75%   81.25%   93.75%
57.33333 66.66667 72.33333 76.33333 80.33333 83.33333 87.33333 91.66667
> ( kmd.p <- with( mDMdat, quantile( rep(P  ,D), qn(nk.p) ) ) )
5.555556% 16.66667% 27.77778% 38.88889%       50% 61.11111% 72.22222% 83.33333% 94.44444%
 1997.667  2000.333  2003.333  2005.667  2007.667  2010.333  2012.333  2014.333  2016.333
> ( kmd.c <- with( mDMdat, quantile( rep(P-A,D), qn(nk.c) ) ) )
   6.25%   18.75%   31.25%   43.75%   56.25%   68.75%   81.25%   93.75%
1912.667 1919.333 1923.667 1927.667 1931.667 1936.333 1942.667 1951.667
> m.md.ap <- glm( D ~ Ns(A,kn=kmd.a) + Ns(P,kn=kmd.p),
+                         offset = log(Y),
+                         family = poisson,
+                           data = subset( mDMdat, sex=="M" ) )
> m.md.aPC <- update( m.md.ap, . ~ . - Ns(P  ,kn=kmd.p) + I(P) +
+                             detrend( Ns(P  ,kn=kmd.p), P  , D ) +
+                             detrend( Ns(P-A,kn=kmd.c), P-A, D ) )
> m.md.apc <- update( m.md.ap, . ~ . + Ns(P-A,kn=kmd.c) )
> m.md.gam <- gam( D ~ s(A) + s(P) + s(C),
+                          offset = log(Y),
```

```
+                          family = poisson,
+                            data = subset( mDMdat, sex=="M" ) )
> f.md.ap  <- update( m.md.ap , data = subset( mDMdat, sex=="W" ) )
> f.md.apc <- update( m.md.apc, data = subset( mDMdat, sex=="W" ) )
> f.md.aPC <- update( m.md.aPC, data = subset( mDMdat, sex=="W" ) )
> f.md.gam <- update( m.md.gam, data = subset( mDMdat, sex=="W" ) )
> Mu.DM[,,"M","ap" ] <- ci.pred( m.md.ap , newdata=nd )[,1]
> Mu.DM[,,"W","ap" ] <- ci.pred( f.md.ap , newdata=nd )[,1]
> Mu.DM[,,"M","apc"] <- ci.pred( m.md.apc, newdata=nd )[,1]
> Mu.DM[,,"W","apc"] <- ci.pred( f.md.apc, newdata=nd )[,1]
> Mu.DM[,,"M","gam"] <- ci.pred( m.md.gam, newdata=nd )[,1]*int
> Mu.DM[,,"W","gam"] <- ci.pred( f.md.gam, newdata=nd )[,1]*int
```

9.5.2 Persons without diabetes

The mortality in the population without diabetes is modeled in exactly the same way as the incidence trends, except we also put in knots in early age too:

```
> ( kmw.a <- with( mnDdat, c( 5, 15,
+                          quantile( rep(A  ,D), qn(nk.a) ) ) ) )
    6.25%    18.75%    31.25%    43.75%    56.25%    68.75%    81.25%    93.75%
 5.00000  15.00000  50.66667  64.33333  71.66667  76.66667  81.33333  85.33333  88.66667  93.66667

> ( kmw.p <- with( mnDdat, quantile( rep(P  ,D), qn(nk.p) ) ) )

5.555556% 16.66667% 27.77778% 38.88889%       50% 61.11111% 72.22222% 83.33333% 94.44444%
 1997.333  1999.333  2001.333  2003.333  2005.667  2008.333  2010.333  2013.333  2015.667

> ( kmw.c <- with( mnDdat, quantile( rep(P-A,D), qn(nk.c) ) ) )

   6.25%   18.75%   31.25%   43.75%   56.25%   68.75%   81.25%   93.75%
1909.667 1916.333 1920.667 1924.667 1929.333 1935.333 1943.333 1956.333

> m.mw.ap <- glm( D ~ Ns(A,kn=kmw.a) + Ns(P,kn=kmw.p),
+                     offset = log(Y),
+                     family = poisson,
+                       data = subset( mnDdat, sex=="M" ) )
> m.mw.aPC <- update( m.mw.ap, . ~ . - Ns(P  ,kn=kmw.p) + I(P) +
+                            detrend( Ns(P  ,kn=kmw.p), P  , D ) +
+                            detrend( Ns(P-A,kn=kmw.c), P-A, D ) )
> m.mw.apc <- update( m.mw.ap, . ~ . + Ns(P-A,kn=kmw.c) )
> m.mw.gam <- gam( D ~ s(A) + s(P) + s(C),
+                          offset = log(Y),
+                          family = poisson,
+                            data = subset( mnDdat, sex=="M" ) )
> f.mw.ap  <- update( m.mw.ap , data = subset( mnDdat, sex=="W" ) )
> f.mw.apc <- update( m.mw.apc, data = subset( mnDdat, sex=="W" ) )
> f.mw.aPC <- update( m.mw.aPC, data = subset( mnDdat, sex=="W" ) )
> f.mw.gam <- update( m.mw.gam, data = subset( mnDdat, sex=="W" ) )
> Mu.nD[,,"M","ap" ] <- ci.pred( m.mw.ap , newdata=nd )[,1]
> Mu.nD[,,"W","ap" ] <- ci.pred( f.mw.ap , newdata=nd )[,1]
> Mu.nD[,,"M","apc"] <- ci.pred( m.mw.apc, newdata=nd )[,1]
> Mu.nD[,,"W","apc"] <- ci.pred( f.mw.apc, newdata=nd )[,1]
> Mu.nD[,,"M","gam"] <- ci.pred( m.mw.gam, newdata=nd )[,1]*int
> Mu.nD[,,"W","gam"] <- ci.pred( f.mw.gam, newdata=nd )[,1]*int
```

9.5.3 Relative mortality (SMR)

The relative mortality in the population without diabetes is modeled in exactly the same way as the mortality and incidence trends:

```
> SMRdat <- rbind( cbind( mDMdat, tp="DM" ),
+                  cbind( mnDdat, tp="nD" ) )
> str( SMRdat )
'data.frame':        354350 obs. of  7 variables:
 $ sex: Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ A  : num  0.667 0.667 0.333 0.667 0.667 ...
 $ P  : num  1997 1999 2000 2000 2000 ...
 $ C  : num  1997 1999 1999 2000 2000 ...
 $ D  : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y  : num  0.1431 0.0226 0.0767 0.3793 0.1758 ...
 $ tp : Factor w/ 2 levels "DM","nD": 1 1 1 1 1 1 1 1 1 1 ...
> ( ksmr.a <- with( SMRdat, c( 5, 15,
+                           quantile( rep(A  ,D), qn(nk.a) ) ) ) )
                     6.25%   18.75%   31.25%   43.75%   56.25%   68.75%   81.25%   93.75%
 5.00000 15.00000 51.66667 64.66667 71.66667 76.66667 81.33333 84.66667 88.66667 93.66667
> ( ksmr.p <- with( SMRdat, quantile( rep(P  ,D), qn(nk.p) ) ) )
5.555556% 16.66667% 27.77778% 38.88889%       50% 61.11111% 72.22222% 83.33333% 94.44444%
 1997.333  1999.333  2001.333  2003.667  2006.333  2008.333  2010.667  2013.333  2015.667
> ( ksmr.c <- with( SMRdat, quantile( rep(P-A,D), qn(nk.c) ) ) )
   6.25%   18.75%   31.25%   43.75%   56.25%   68.75%   81.25%   93.75%
1910.333 1916.667 1921.333 1925.333 1929.667 1935.333 1943.333 1955.333
> m.smr.ap <- glm( D ~ Ns(A,kn=ksmr.a,int=TRUE):tp + I(P):tp,
+                     offset = log(Y),
+                     family = poisson,
+                       data = subset( SMRdat, sex=="M" ) )
> #m.smr.gam <- gam( D ~ s(A,by=tp) + I(P):tp,
> #                        offset = log(Y),
> #                        family = poisson,
> #                          data = subset( SMRdat, sex=="M" ) )
> f.smr.ap  <- update( m.smr.ap , data = subset( SMRdat, sex=="W" ) )
> #f.smr.gam <- update( m.smr.gam, data = subset( SMRdat, sex=="W" ) )
```

We can show how SMR varies by calendar time:

```
> round((ci.exp(m.smr.ap,subset="P",ctr.mat=rbind(c(1,-1)))-1)*100,1)[c(1,3,2)]
[1] -1.1 -1.0 -1.2
> round((ci.exp(f.smr.ap,subset="P",ctr.mat=rbind(c(1,-1)))-1)*100,1)[c(1,3,2)]
[1] -1.2 -1.0 -1.3
> nx <- data.frame( A=30:90, P=2010, tp="DM" )
> nr <- data.frame( A=30:90, P=2010, tp="nD" )
> SMRma <- ci.exp( m.smr.ap, list(nx,nr) )
> SMRfa <- ci.exp( f.smr.ap, list(nx,nr) )
> matshade( nx$A, cbind( SMRma, SMRfa ), plot=TRUE,
+           log="y",  col=c(4,2), lwd=2, ylim=c(0.9,6),
+           xlab="Age", ylab="SMR ( DM vs. no DM) at 2010" )
> abline( h=1, lty=3 )
> abline( h=3:12/2, col=gray(0.95) )
> axis( side=2, at=3:12/2, labels=NA, tcl=-0.3 )
```

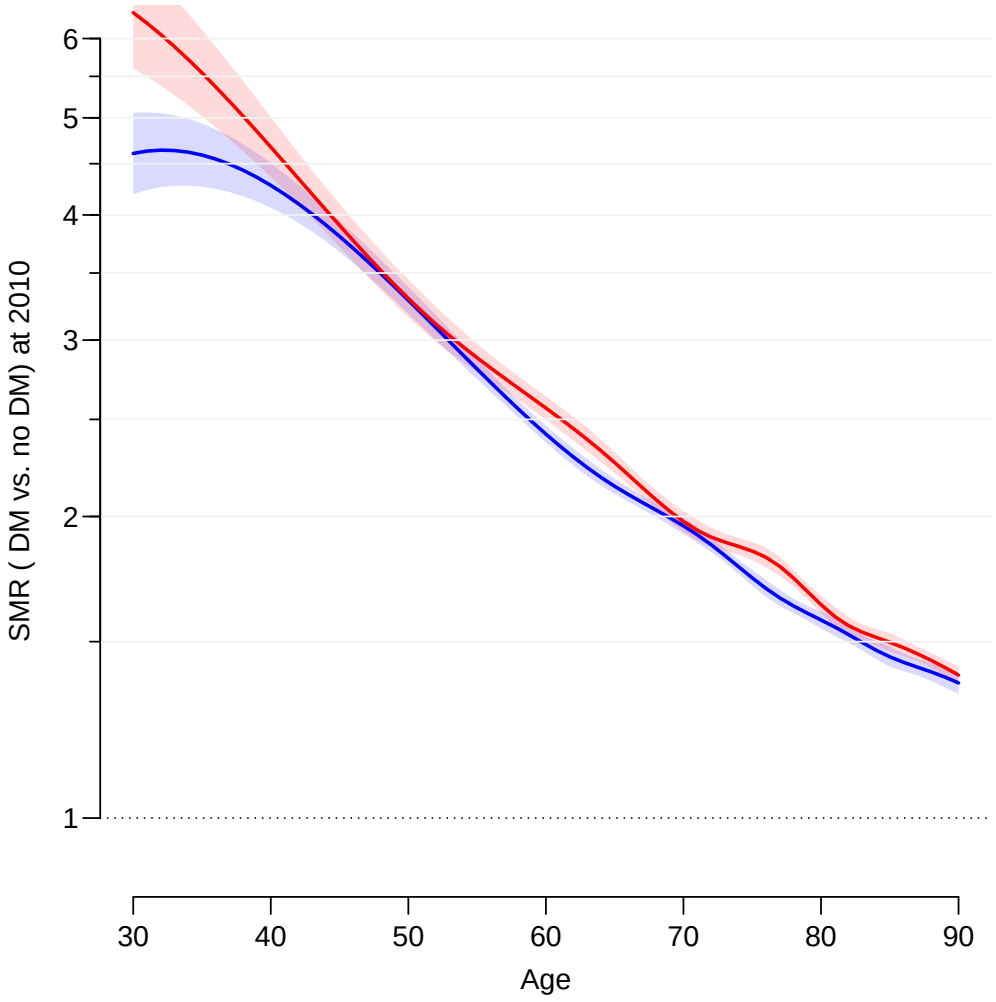


Figure 9.6: *SMR between DM and no-DM.* `./graph/pr-rates-ageSMR`

9.6 Average trends

The average annual trends in all of the rates (in %) can now be summarized:

```
> mort.chg <- (rbind( ci.exp(m.md.aPC,subset="I\\(P"),
+                     ci.exp(f.md.aPC,subset="I\\(P"),
+                     ci.exp(m.mw.aPC,subset="I\\(P"),
+                     ci.exp(f.mw.aPC,subset="I\\(P"),
+                     ci.exp(m.smr.ap,subset="P",ctr.mat=rbind(c(1,-1))),
+                     ci.exp(f.smr.ap,subset="P",ctr.mat=rbind(c(1,-1))) )-1)*100
> rownames( mort.chg ) <- c("Mortality change,   DM: Men",
+                           "                        Women",
+                           "Mortality change, noDM: Men",
+                           "                        Women",
+                           "SMR change:               Men",
+                           "                        Women")
> round( rbind( inc.chg, mort.chg ), 2 )
                              exp(Est.)  2.5% 97.5%
DM incidence change Men total:      2.95  2.82  3.09
                     pre-2011       4.18  4.07  4.30
                    post-2011      -4.45 -4.75 -4.16
                Women total:       2.79  2.64  2.93
                     pre-2011      3.79  3.67  3.92
                    post-2011      -4.67 -5.00 -4.34
Mortality change,   DM: Men        -3.93 -4.04 -3.82
                      Women        -3.48 -3.61 -3.36
Mortality change, noDM: Men        -2.89 -2.94 -2.84
                      Women        -2.46 -2.51 -2.41
SMR change:             Men        -1.11 -1.22 -0.99
                      Women        -1.16 -1.28 -1.03
> # halving/doubing time of rates
> round( log(0.5)/log(1-abs(rbind(inc.chg,mort.chg)/100)), 1 )
                              exp(Est.) 2.5% 97.5%
DM incidence change Men total:      23.1 24.3  22.1
                     pre-2011       16.2 16.7  15.8
                    post-2011       15.2 14.3  16.3
                Women total:       24.5 25.9  23.3
                     pre-2011      17.9 18.6  17.3
                    post-2011      14.5 13.5  15.6
Mortality change,   DM: Men         17.3 16.8  17.8
                      Women         19.5 18.9  20.3
Mortality change, noDM: Men         23.7 23.3  24.1
                      Women         27.8 27.3  28.4
SMR change:             Men         62.3 56.2  69.8
                      Women         59.5 53.6  66.9
```

Thus it appears that the incidence rates of diabetes overall are increasing by some 2.5% per year, while mortality rates are decreasing 3.5% per year for persons with diabetes, but only 2.5% per year for persons without — the latter two slightly more for men than for women.

9.7 Time trends in rates

In order to show how the models predict the time trends in diabetes incidence and mortality, we make a graphical display of the estimated rates in ages 20,. . . ,90 versus calendar time, and of the estimated rates at the beginning of 1996, 2000, . . . , 2016 versus age.

9.7.1 Age-specific rates by date

A brief overview of the mortality and incidence rates over time. Recall that for the sake of future calculations the rates in the arrays are in units of $1/\texttt{int} = 1/12$; hence if we want to have rates not in units of `int` (person-years) but in units of 1000 person-years we must multiply by $1000/\texttt{int}$. We also define some hues of blue and red for use in the graph:

```
> ytl <- function(x) x[x>10^par("usr")[3] & x<10^par("usr")[4]]
> gsc <- c("#11CC11","#66CC66","#99CC99")
> rsc <- function(n) rgb( cbind(       255    , seq(0,200,,n), seq(0,200,,n) ), max=255 )
> bsc <- function(n) rgb( cbind( seq(0,200,,n), seq(0,200,,n),       255     ), max=255 )
> pts <- as.numeric( dimnames(Lambda)[[2]] )[1:252]
> ( dimnames(Lambda)[[1]][agr <- seq(240,1080,120)] )
[1] "19.9583333333333" "29.9583333333333" "39.9583333333333" "49.9583333333333"
[5] "59.9583333333333" "69.9583333333333" "79.9583333333333" "89.9583333333333"
> yticks <- outer( c(1:9), 10^(-3:3) )
> rpl <-
+ function( Lambda, sx, yl )
+ {
+ plot( NA, ylim=c(0.03,200)*2, xlim=c(1996,2017), yaxt="n", xaxt="n",
+           ylab=yl, yaxs="i", xlab="", log="y" )
+ abline( h=yticks, col=gray(0.8) )
+ matlines( pts, t(Lambda[agr,1:252,sx,"apc"])*1000/int,
+          lty=1, lwd=3, type="l",
+          col=if(sx=="M") bsc(8) else rsc(8) )
+ mtext( side=2, yl, line=2.5, las=0, cex=0.66 )
+ }
> par( mfrow=c(3,2), mar=c(1,1,1,1), oma=c(3,3,0,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> rpl( Lambda, "M", "DM incidence per 1000 PY" )
>                          axis( side=2 ) ; axis( side=2, at=ytl(yticks), labels=NA )
> text( 1996, 200, "Men", adj=c(0,1), cex=1.5, col="blue" )
> rpl( Lambda, "W", "" )
> text( 1996, 200, "Women", adj=c(0,1), cex=1.5, col="red" )
> rpl(  Mu.nD, "M", "Population mortality per 1000 PY" )
>                          axis( side=2 ) ; axis( side=2, at=ytl(yticks), labels=NA )
> rpl(  Mu.nD, "W", "" )
> rpl(  Mu.DM, "M", "DM mortality per 1000 PY" )
>                          axis( side=2 ) ; axis( side=2, at=ytl(yticks), labels=NA )
>                          axis( side=1 ) ; axis( side=1, at=1996:2017, labels=NA, tcl=-0.3 )
> rpl(  Mu.DM, "W", "" ) ; axis( side=1 ) ; axis( side=1, at=1996:2017, labels=NA, tcl=-0.3 )
> mtext( "Date of follow-up", side=1, line=2, outer=TRUE, cex=0.66 )
```

Illustrative plots

These are plots where more and more rate curves are added successively aimed at inclusion in a presentation:

```
> rplr <-
+ function( Lambda, agr, sx, yl )
+ {
+ plot( NA, ylim=c(0.3,200), xlim=c(1996,2017), # yaxt="n", xaxt="n",
+           ylab=yl, xlab="", log="y" )
+ abline( h=yticks, col=gray(0.8) )
+ wh <- match(agr,2:9*10)
+ cl <- if(sx=="M") bsc(8)[wh] else rsc(8)[wh]
```

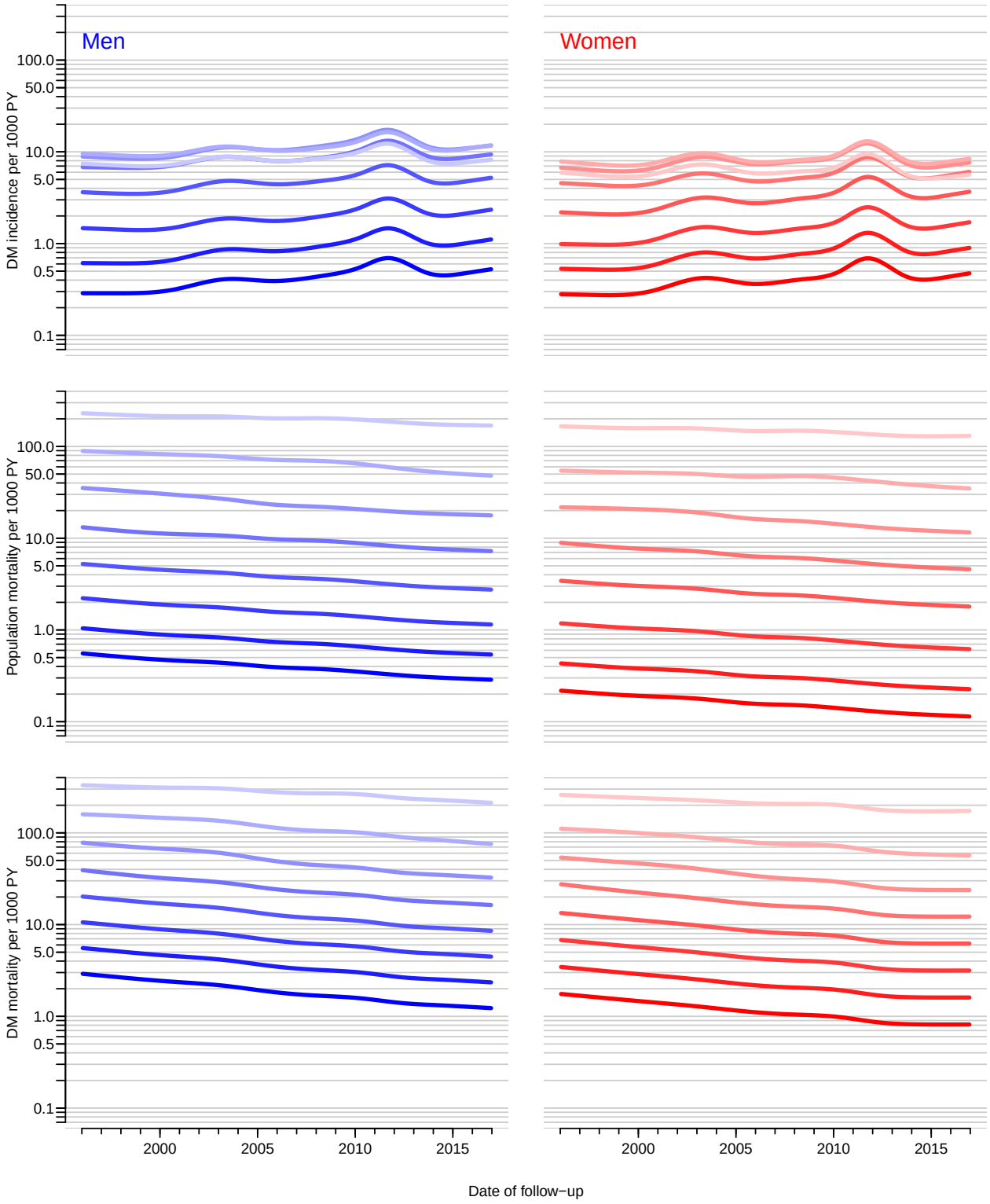


Figure 9.7: *Trends in incidence and mortality rates at ages 20 (darkest), 30, . . . , 90 (lightest), as estimated from separate age-period-cohort models for men and women.* `./graph/pr-rates-ratesbyper`

```
+ pts <- as.numeric(dimnames(Lambda)[[2]][1:252])
+ matlines( pts, t(Lambda[agr*12+1,1:252,sx,"apc"])*1000/int,
+           lty=1, lwd=3, type="l", col=cl )
+ text( 2016, Lambda[agr*12+1,252,sx,"apc"]*1000/int, agr, adj=0, col=cl )
+ }
> rplr( Lambda, c(2,2)*10, "M", "DM incidence per 1000 PY")
```

9.7.2 Date specific rates by age

First we take a look at the predicted rates in the arrays and tease out the indices on the date-scale that we shall use for plotting the rate by age every 4 years:

```
> str( Lambda )
 num [1:1200, 1:528, 1:2, 1:9] 6.14e-06 6.18e-06 6.22e-06 6.26e-06 6.30e-06 ...
 - attr(*, "dimnames")=List of 4
  ..$ a  : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" ...
  ..$ p  : chr [1:528] "1996.04166666667" "1996.125" "1996.20833333333" "1996.29166666667" ...
  ..$ sex: chr [1:2] "M" "W"
  ..$ mod: chr [1:9] "ap" "apc" "gam" "LCa" ...
> range( as.numeric(dimnames(Lambda)[[2]]) )
[1] 1996.042 2039.958
> wh <- 1+c(0,2,4,6,8,10)*24
> as.numeric(dimnames(Lambda)[[2]])[wh]
[1] 1996.042 2000.042 2004.042 2008.042 2012.042 2016.042
> as.numeric(dimnames(Mu.DM)[[2]])[wh]
[1] 1996.042 2000.042 2004.042 2008.042 2012.042 2016.042
> as.numeric(dimnames(Mu.nD)[[2]])[wh]
[1] 1996.042 2000.042 2004.042 2008.042 2012.042 2016.042
```

We now plot the rates by age for each of these calendar years, as fitted by the APC model:

```
> a.pt <- as.numeric(dimnames(Lambda)[[1]])
> a.pt <- ndn(Lambda,1)
> rtpl <-
+ function( Lambda, ylb, clr, sx, md, leg=FALSE )
+ {
+ plot( NA, ylim=c(0.3,200), xlim=c(20,95), yaxt="n", xaxt="n",
+           ylab=ylb, xlab="", log="y" )
+ abline( h=yticks, col=gray(0.8) )
+ matlines( a.pt, Lambda[,wh,sx,md]*(1000/int),
+           type="l", lty=1, lwd=3, col=clr )
+ if(sx=="M") axis(side=2)
+ if(leg)
+ text( 20, 200*0.6^(1:length(wh)), round(ndn(Lambda,2)[wh]), col=clr, adj=0 )
+ }
> for( md in c("ap","apc") )
+ par( mfcol=c(3,2), mar=c(1,1,1,1), oma=c(3,3,0,0), mgp=c(3,1,0)/1.6,
+      las=1, bty="n" )
> for( sx in c("M","W") )
+     {
+ lw <- length(wh)
+ clr <- if(sx=="M") bsc(lw)[lw:1] else rsc(lw)[lw:1]
+ clr[5] <- gray(0.5)
```

```
+ rtpl( Lambda, if(sx=="M") "DM incidence per 1000 PY" else "", clr, sx, md, leg=TRUE )
+ rtpl( Mu.nD, if(sx=="M") "Population mortality per 1000 PY" else "", clr, sx, md )
+ rtpl( Mu.DM, if(sx=="M") "DM mortality per 1000 PY" else "", clr, sx, md )
+ axis( side=1 )
+ axis(side=1,at=seq(20,95,10),tcl=-0.5,labels=NA)
+ axis(side=1,at=seq(20,95, 5),tcl=-0.3,labels=NA)
+     }
> mtext( "Age at follow-up", side=1, line=2, outer=TRUE, cex=0.66 )
```

9.8 Extrapolation of rates

It is possible to extrapolate the rates beyond the observed dates by simply extending the linear part of the natural splines; in fact this is already done in the rate-objects `Lambda`, `Mu.nD` and `Mu.DM`. However, as seen in figure 9.7 the predicted decline in diabetes rates is presumably way too dramatic.

We therefore repeat the attenuation exercise for mortality rates too, but we shall not add future increases to the rates as we do for sensitivity analyses of the incidence rates.

```
> str( Lambda )
 num [1:1200, 1:528, 1:2, 1:9] 6.14e-06 6.18e-06 6.22e-06 6.26e-06 6.30e-06 ...
 - attr(*, "dimnames")=List of 4
  ..$ a  : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" ...
  ..$ p  : chr [1:528] "1996.04166666667" "1996.125" "1996.20833333333" "1996.29166666667" ...
  ..$ sex: chr [1:2] "M" "W"
  ..$ mod: chr [1:9] "ap" "apc" "gam" "LCa" ...

> args( damp )

function (f, t, h, delta = 0, ell = 0, logf = FALSE)
NULL

> # where is the future ?
> wh.p <- (1:length(p.pt))[p.pt>=2017]
> # and what are the dates we are considering ?
> p.pt <- p.pt[p.pt>=2017]
> range( p.pt )

[1] 2017.042 2039.958

> dimnames(Lambda)[[4]]

[1] "ap"  "apc" "gam" "LCa" "att" "fix" "p20" "p40" "p60"
```

Digression on `apply`

Note that when we apply a function that returns a vector to an array the returned vector will form the first dimension in the result

```
> ( M <- matrix( 1:6, 3, 2 ) )
     [,1] [,2]
[1,]    1    4
[2,]    2    5
[3,]    3    6
```

The second argument to `apply` is a vector of dimensions of the object (the first argument) to keep, so the returned result of `apply` will have the function result as the first dimension, and the dimensions mentioned in the second argument as the following dimensions, as these two examples show:

```
> apply( M, 1, rev )
     [,1] [,2] [,3]
[1,]    4    5    6
[2,]    1    2    3
> apply( M, 1, function(x) c(x,rev(x)) )
     [,1] [,2] [,3]
[1,]    1    2    3
[2,]    4    5    6
[3,]    4    5    6
[4,]    1    2    3
```

This is what we shall do here, hence the `t()` in the following (and in the code below where predictions are altered):

```
> t( apply( M, 1, rev ) )
     [,1] [,2]
[1,]    4    1
[2,]    5    2
[3,]    6    3
```

. . . end of digression.

9.8.1 Attenuated extrapolations

Here we make the changes to the rates in the prediction period indicated by `wh.p`; we make a fix version of all rates, that just extends rates at the last time point. Attenuated versions are made where we attenuate the rate slopes by a halving time of 4 years, and finally for the incidence rates we make versiosn where we have a very short halving time, but add fixed increases of 2, 4 and 6 percent per year.

```
> wh.p[0:2]
[1] 253 254
> dimnames( Lambda )[[2]][wh.p[0:2]]
[1] "2017.04166666667" "2017.125"
> system.time(
+ for( is in dimnames(Lambda)[['sex']] )
+    {
+ Lambda[,wh.p,is,"fix"] <- t(apply(Lambda[,wh.p,is,"apc"],1,damp,t=p.pt,h=0.0,delta=0.0))
+ Lambda[,wh.p,is,"att"] <- t(apply(Lambda[,wh.p,is,"apc"],1,damp,t=p.pt,h=5.0,delta=0.0))
+ Lambda[,wh.p,is,"p20"] <- t(apply(Lambda[,wh.p,is,"apc"],1,damp,t=p.pt,h=1.0,delta=2.0,
+                                                                          ell=2))
+ Lambda[,wh.p,is,"p40"] <- t(apply(Lambda[,wh.p,is,"apc"],1,damp,t=p.pt,h=1.0,delta=4.0,
+                                                                          ell=2))
+ Lambda[,wh.p,is,"p60"] <- t(apply(Lambda[,wh.p,is,"apc"],1,damp,t=p.pt,h=1.0,delta=6.0,
+                                                                          ell=2))
+  Mu.DM[,wh.p,is,"fix"] <- t(apply( Mu.DM[,wh.p,is,"apc"],1,damp,t=p.pt,h=0.0,delta=0.0))
+  Mu.DM[,wh.p,is,"att"] <- t(apply( Mu.DM[,wh.p,is,"apc"],1,damp,t=p.pt,h=5.0,delta=0.0))
+  Mu.nD[,wh.p,is,"fix"] <- t(apply( Mu.nD[,wh.p,is,"apc"],1,damp,t=p.pt,h=0.0,delta=0.0))
+  Mu.nD[,wh.p,is,"att"] <- t(apply( Mu.nD[,wh.p,is,"apc"],1,damp,t=p.pt,h=5.0,delta=0.0))
+    } )
```

```
   user  system elapsed
  16.97    0.12   17.13
```

In principle we could use the same attenuation along the age-dimension within each cohort, but we are only interested in the calendar time, and even though the cohort-specific curves increase *both* by age and calendar time, we are not interested in attenuation of the *combined effect* of age and calendar time. Or put differently — we are not interested in any attenuation along the age-scale.

For the sake of easier use we fill in the prediction dimenasions for the first years be for we save:

```
> str( Lambda )
 num [1:1200, 1:528, 1:2, 1:9] 6.14e-06 6.18e-06 6.22e-06 6.26e-06 6.30e-06 ...
 - attr(*, "dimnames")=List of 4
  ..$ a  : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" ...
  ..$ p  : chr [1:528] "1996.04166666667" "1996.125" "1996.20833333333" "1996.29166666667" ...
  ..$ sex: chr [1:2] "M" "W"
  ..$ mod: chr [1:9] "ap" "apc" "gam" "LCa" ...
> table( is.na(Lambda[700,,1,"att"]))
FALSE  TRUE
  276   252
> str(Mu.DM)
 num [1:1200, 1:528, 1:2, 1:9] 6.22e-05 6.26e-05 6.29e-05 6.33e-05 6.36e-05 ...
 - attr(*, "dimnames")=List of 4
  ..$ a  : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" ...
  ..$ p  : chr [1:528] "1996.04166666667" "1996.125" "1996.20833333333" "1996.29166666667" ...
  ..$ sex: chr [1:2] "M" "W"
  ..$ mod: chr [1:9] "ap" "apc" "gam" "LCa" ...
> for( i in 5:9 ) Lambda[,1:252,,i] <- Lambda[,1:252,,"apc"]
> for( i in 5:6 )
+    {
+  Mu.DM[,1:252,,i] <-  Mu.DM[,1:252,,"apc"]
+  Mu.nD[,1:252,,i] <-  Mu.nD[,1:252,,"apc"]
+    }
> save( Lambda, Mu.DM, Mu.nD, file="../nydata/allrates.Rda" )
```

9.9 The projected rates beyond 2016

```
> load( file="../nydata/allrates.Rda" )
```

Now we have the three types of rates in the illness-death model in the arrays `Lambda`, `Mu.DM` and `Mu.nD`, so we plot these for select ages as a function of time — again recall they are in units of events per `int`, so we rescale as before to get rates per 1000 PY:

```
> pts <- as.numeric( dimnames(Lambda)[[2]] )
> rpl <-
+ function( Lambda, sx, yl=NULL, xm=2040, inc=NULL, icol=NULL, ylm=c(0.06,400) )
+ {
+ clr <- if( sx=="M") bsc(8) else rsc(8)
+ plot( NA, log="y", yaxt="n", yaxs="i", ylim=ylm, xlim=c(1996,xm),
```

```
+       ylab="", xaxt="n" )
+ abline( h=yticks, col=gray(0.8) )
+ matlines( pts, t(Lambda[agr,,sx,"apc"])*1000/int,
+           lty=1, lwd=4, type="l", col=clr )
+ matlines( pts, t(Lambda[agr,,sx,"att"])*1000/int,
+           lty=1, lwd=2, type="l", col="forestgreen" )
+ matlines( pts, t(Lambda[agr,,sx,"fix"])*1000/int,
+           lty=1, lwd=1, type="l", col="black" )
+ if( !is.null(inc) )
+ for( ii in 1:length(inc) )
+    matlines( pts, t(Lambda[agr,,sx,inc[ii]])*1000/int,
+              lty=1, lwd=1, type="l", col=icol[ii] )
+ abline( v=2017, lty=3 )
+ mtext( side=2, yl, line=2.5, las=0, cex=0.66 )
+ }
> do.pl <- function( xm=2030 )
+     {
+ rpl( Lambda, "M", "DM incidence per 1000 PY", xm=xm, inc=c("p20","p40","p60"), icol=gsc )
+ axis( side=2 ) ; axis( side=2, at=ytl(yticks), labels=NA, tcl=-0.2 )
+ text( 1996, 380, "Men", adj=c(0,1), cex=1.5, col="blue" )
+ rpl( Lambda, "W", "", xm=xm, inc=c("p20","p40","p60"), icol=gsc )
+ text( 1996, 380, "Women", adj=c(0,1), cex=1.5, col="red" )
+ rpl(  Mu.nD, "M", "Mortality (no DM) per 1000 PY", xm=xm )
+ axis( side=2 ) ; axis( side=2, at=ytl(yticks), labels=NA, tcl=-0.2 )
+ rpl(  Mu.nD, "W", "", xm=xm )
+ rpl(  Mu.DM, "M", "DM mortality per 1000 P", xm=xm )
+ axis( side=2 ) ; axis( side=2, at=ytl(yticks), labels=NA, tcl=-0.2 )
+ axis( side=1 ) ; axis( side=1, at=seq(1995,xm,5), labels=NA, tcl=-0.4 )
+                  axis( side=1, at=1996:xm, labels=NA, tcl=-0.2 )
+ rpl(  Mu.DM, "W", "", xm=xm )
+ axis( side=1 )
+ axis( side=1, at=seq(1995,xm,5), labels=NA, tcl=-0.4 )
+ axis( side=1, at=seq(1995,xm,1), labels=NA, tcl=-0.2 )
+ mtext( "Date of follow-up", side=1, line=2, outer=TRUE, cex=0.66 )
+     }
> par( mfrow=c(3,2), mar=c(1,1,1,1), oma=c(3,3,0,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> do.pl( xm=2040 )


> par( mfrow=c(3,2), mar=c(1,1,1,1), oma=c(3,3,0,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> do.pl( xm=2030 )


> par( mfcol=c(1,1), mar=c(1,1,1,1), oma=c(3,3,0,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> xm <- 2030
> rpl( Lambda, "M", "", xm=xm, inc=c("p20","p40","p60"), icol=gsc, ylm=c(0.2,40) )
> axis( side=2 ) ; axis( side=2, at=ytl(yticks), labels=NA, tcl=-0.2 )
> axis( side=1 ) ; axis( side=1, at=seq(1995,xm,5), labels=NA, tcl=-0.4 )
>                  axis( side=1, at=1996:xm, labels=NA, tcl=-0.2 )
> mtext( "Date of follow-up", side=1, line=2, outer=TRUE )
> mtext( rep("DM incidence rates per 1000 PY",2),
+        side=2, line=2, outer=TRUE, las=0 )


> par( mfcol=c(2,3), mar=c(1,1,1,1), oma=c(3,3,0,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> do.pl <- function( xm=2030 )
+     {
+ rpl( Lambda, "M", "", xm=xm, inc=c("p20","p40","p60"), icol=gsc )
```

```
+ axis( side=2 ) ; axis( side=2, at=ytl(yticks), labels=NA, tcl=-0.2 )
+ text( 1996, 380, "Men", adj=c(0,1), cex=1.5, col="blue" )
+ text( xm, 380, "DM incidence", adj=c(1,1), cex=1.5 )
+ rpl( Lambda, "W", "", xm=xm, inc=c("p20","p40","p60"), icol=gsc )
+ axis( side=2 ) ; axis( side=2, at=ytl(yticks), labels=NA, tcl=-0.2 )
+ axis( side=1 ) ; axis( side=1, at=seq(1995,xm,5), labels=NA, tcl=-0.4 )
+                   axis( side=1, at=1996:xm, labels=NA, tcl=-0.2 )
+ text( 1996, 380, "Women", adj=c(0,1), cex=1.5, col="red" )
+ rpl(  Mu.nD, "M", "", xm=xm )
+ text( xm, 380, "Mortality (no DM)", adj=c(1,1), cex=1.5 )
+ rpl(  Mu.nD, "W", "", xm=xm )
+ axis( side=1 ) ; axis( side=1, at=seq(1995,xm,5), labels=NA, tcl=-0.4 )
+                   axis( side=1, at=1996:xm, labels=NA, tcl=-0.2 )
+ rpl(  Mu.DM, "M", "", xm=xm )
+ text( xm, 380, "DM mortality", adj=c(1,1), cex=1.5 )
+ rpl(  Mu.DM, "W", "", xm=xm )
+ axis( side=1 )
+ axis( side=1, at=seq(1995,xm,5), labels=NA, tcl=-0.4 )
+ axis( side=1, at=seq(1995,xm,1), labels=NA, tcl=-0.2 )
+ mtext( "Date of follow-up", side=1, line=2, outer=TRUE, cex=0.66 )
+ mtext( rep("Rates per 1000 PY",2), at=c(1,3)/4, side=2, line=2, outer=TRUE, cex=0.66, las=0
+      }
> do.pl( xm=2030 )
```

Here is a piece of code for illustration of the incidence projection:

```
> rpl1 <-
+ function( Lambda, sx, yl=NULL, inc=NULL, icol=NULL )
+ {
+ clr <- if( sx=="M") bsc(8) else rsc(8)
+ par( mar=c(3,4,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ plot( NA, log="y",
+       ylab="", # yaxt="n", xaxt="n",
+       ylim=c(0.3,50), xlim=c(1996,2030) )
+ abline( h=yticks, col=gray(0.8) )
+ matlines( pts, t(Lambda[agr,,sx,"apc"])*1000/int,
+           lty=1, lwd=5, type="l", col=clr )
+ if( !is.null(inc) )
+ for( ii in 1:length(inc) )
+    matlines( pts, t(Lambda[agr,,sx,inc[ii]])*1000/int,
+              lty=1, lwd=2, type="l", col=icol[ii] )
+ abline( v=2017, lty=3 )
+ mtext( side=2, yl, line=2.5, las=0 )
+ }
> par( mfrow=c(3,2), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> rpl1( Lambda, "M", inc=NULL, yl="Diabetes incidence rates / 1000 PY" )
> text( 1996, 35, "Incidence rates projected by APC model", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "M", inc="fix", icol="black" )
> text( 1996, 35, "Incidence rates fixed at 2016 level", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "M", inc="att", icol="forestgreen", yl="Diabetes incidence rates / 1000 PY" )
> text( 1996, 35, "Incidence rate increase halved every 4 years", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "M", inc="p20", icol="forestgreen" )
> text( 1996, 35, "Incidence rates increasing 2%/y from 2016 level", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "M", inc="p40", icol="forestgreen", yl="Diabetes incidence rates / 1000 PY" )
> text( 1996, 35, "Incidence rates increasing 4%/y from 2016 level", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "M", inc="p60", icol="forestgreen" )
> text( 1996, 35, "Incidence rates increasing 6%/y from 2016 level", adj=c(0,0), cex=1.3 )
```

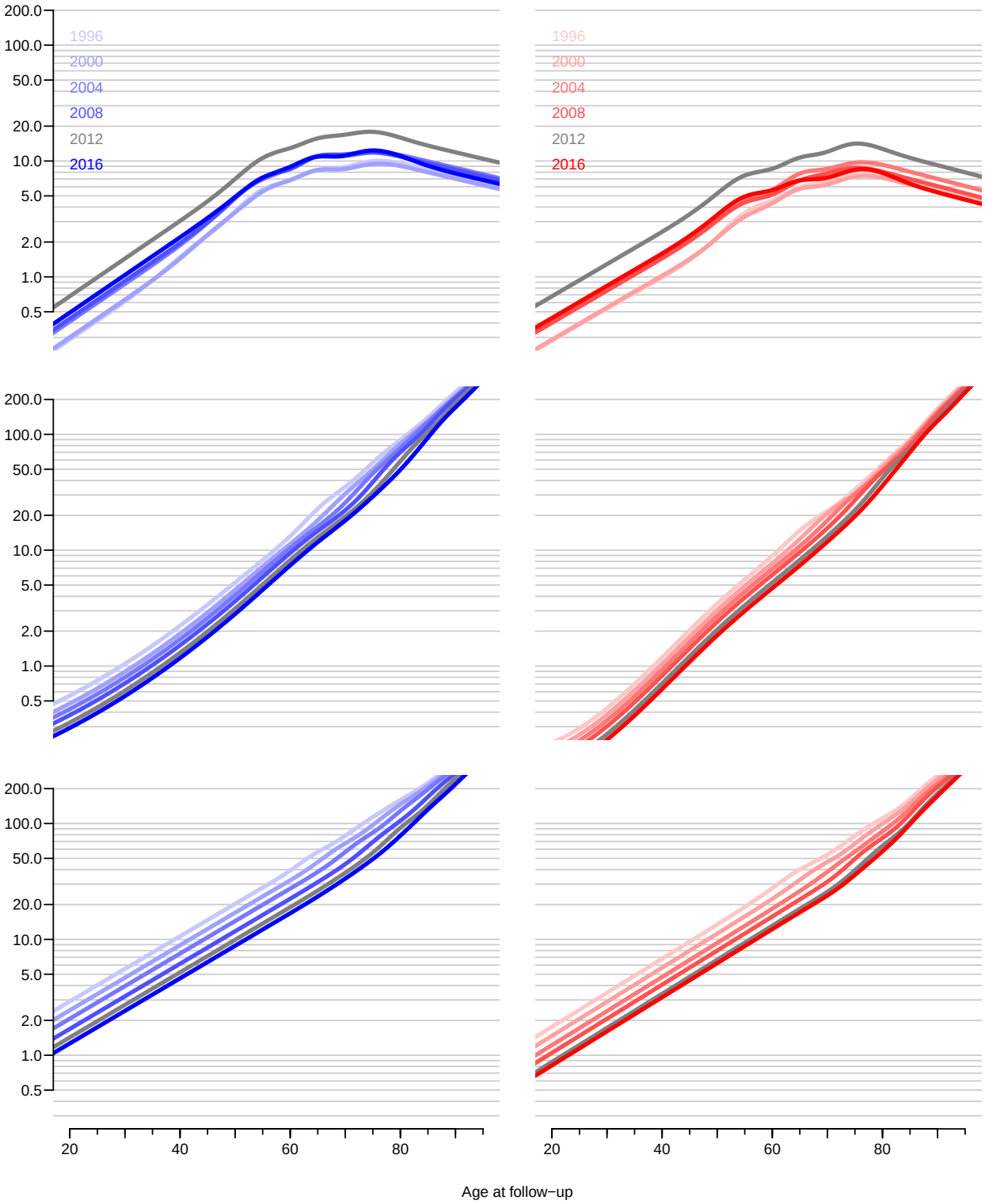


Figure 9.8: *Incidence and mortality rates by age from 1996 through 2016, as estimated from separate age-period-cohort models for men and women.* `./graph/pr-rates-ratesbyage`

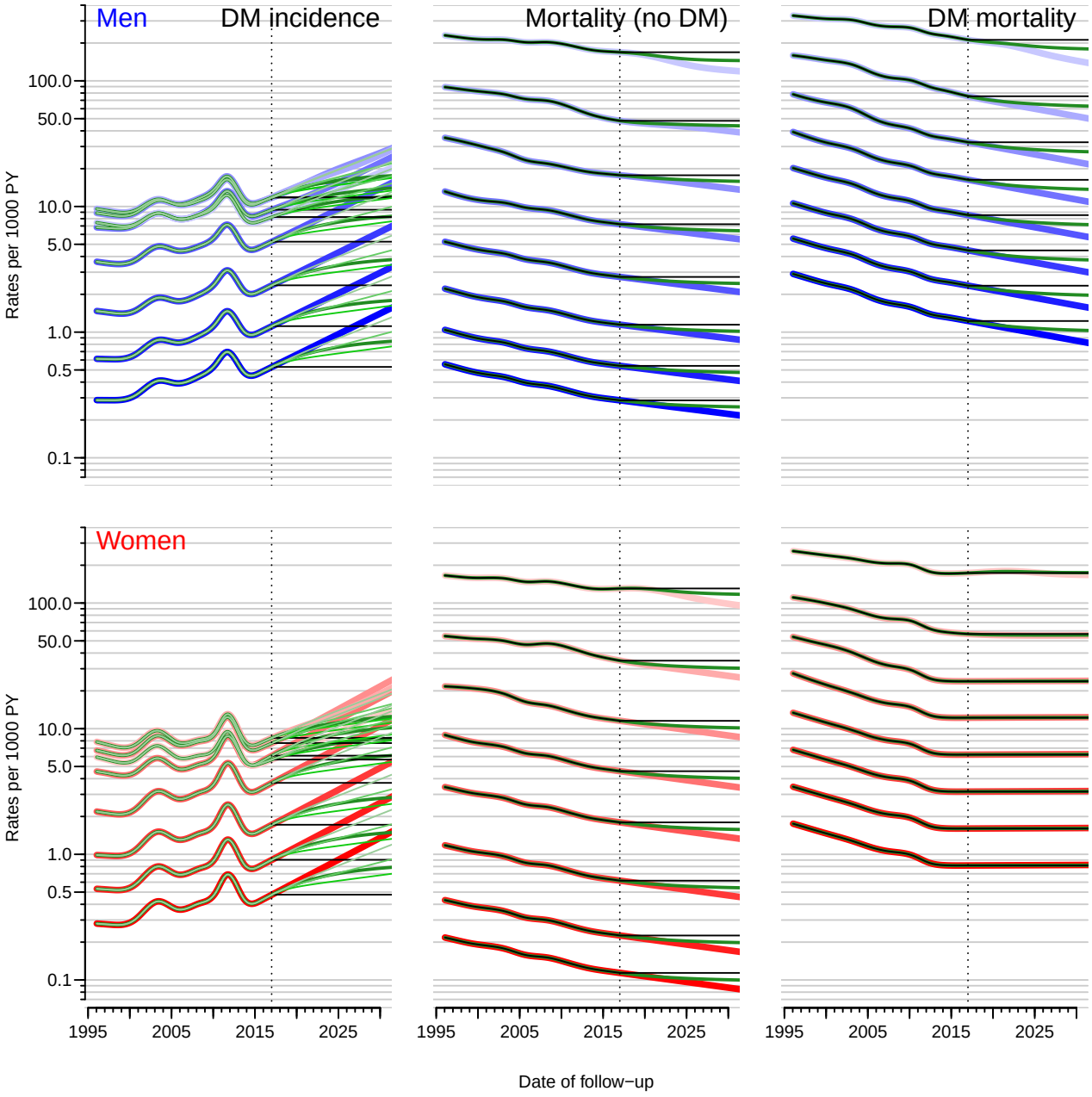


Figure 9.9: *Trends 1996–2016 and projections 2017–2030 of incidence and mortality rates for ages 20 (darkest), 30, . . . , 90 (lightest), as estimated from the age-period-cohort models and projcted by naive extrapolation of the natural splines.*
The thin black lines indicate predictions fixing all rates at the level at the end of 2016, the thick green lines using an attenuation of the trend halving the slope every 4 years, and the thin green lines for incidence rates represent an annual increase of 2.0, 4.0 and 6.0 % per year (dark to light) added (phased in over 1 year).
The vertical dotted lines indicate the end of available data. `./graph/pr-rates-projrates30h`

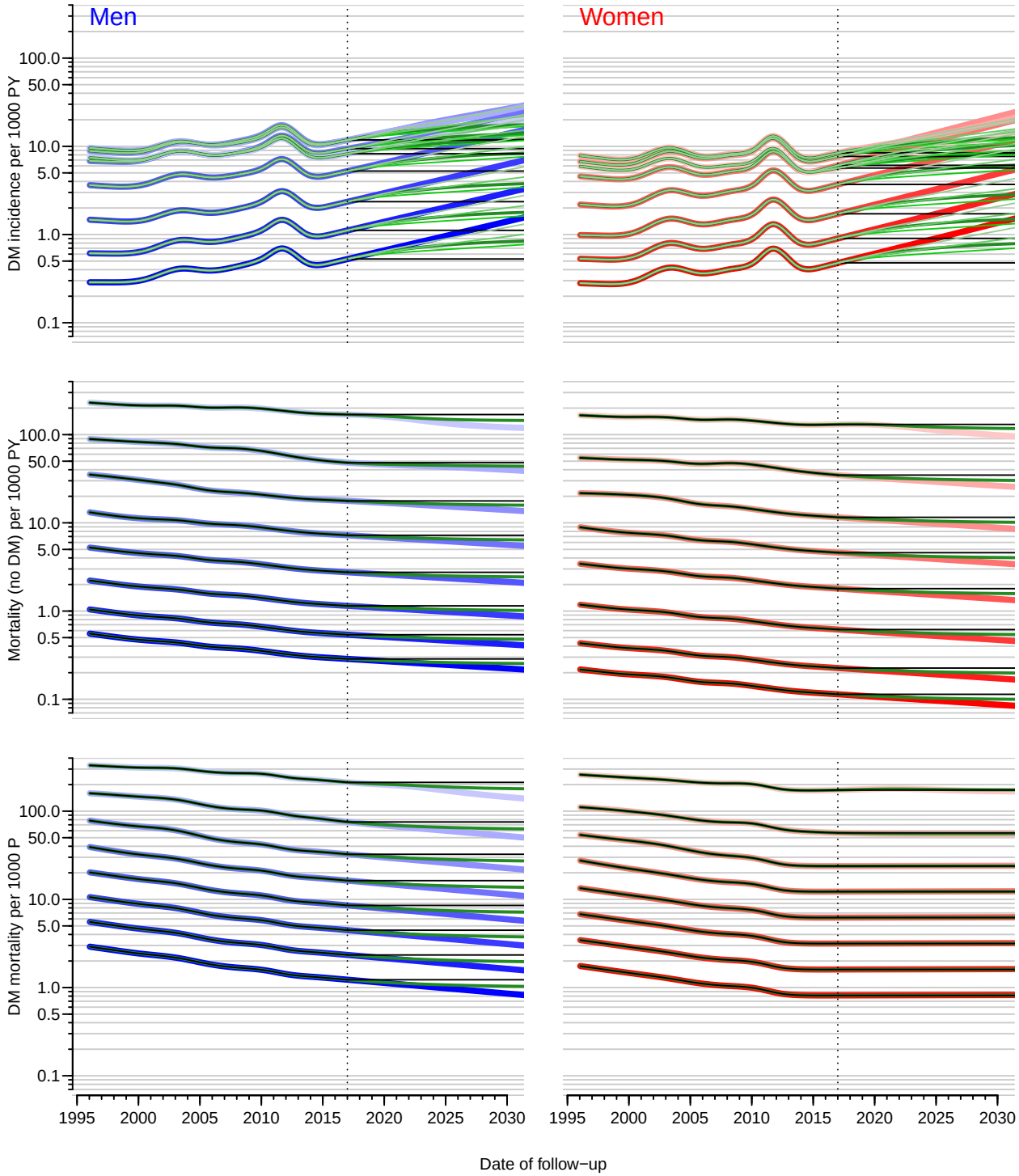


Figure 9.10: *Trends 1996–2016 and projections 2017–2030 of incidence and mortality rates for ages 20 (darkest), 30, . . . , 90 (lightest), as estimated from the age-period-cohort models and projcted by naive extrapolation of the natural splines.*
The thin black lines indicate predictions fixing all rates at the level at the end of 2016, the thick green lines using an attenuation of the trend halving the slope every 4 years, and the thin green lines for incidence rates represent an annual increase of 2.0, 4.0 and 6.0 % per year (dark to light) added (phased in over 1 year).
The vertical dotted lines indicate the end of available data. `./graph/pr-rates-projrates30`

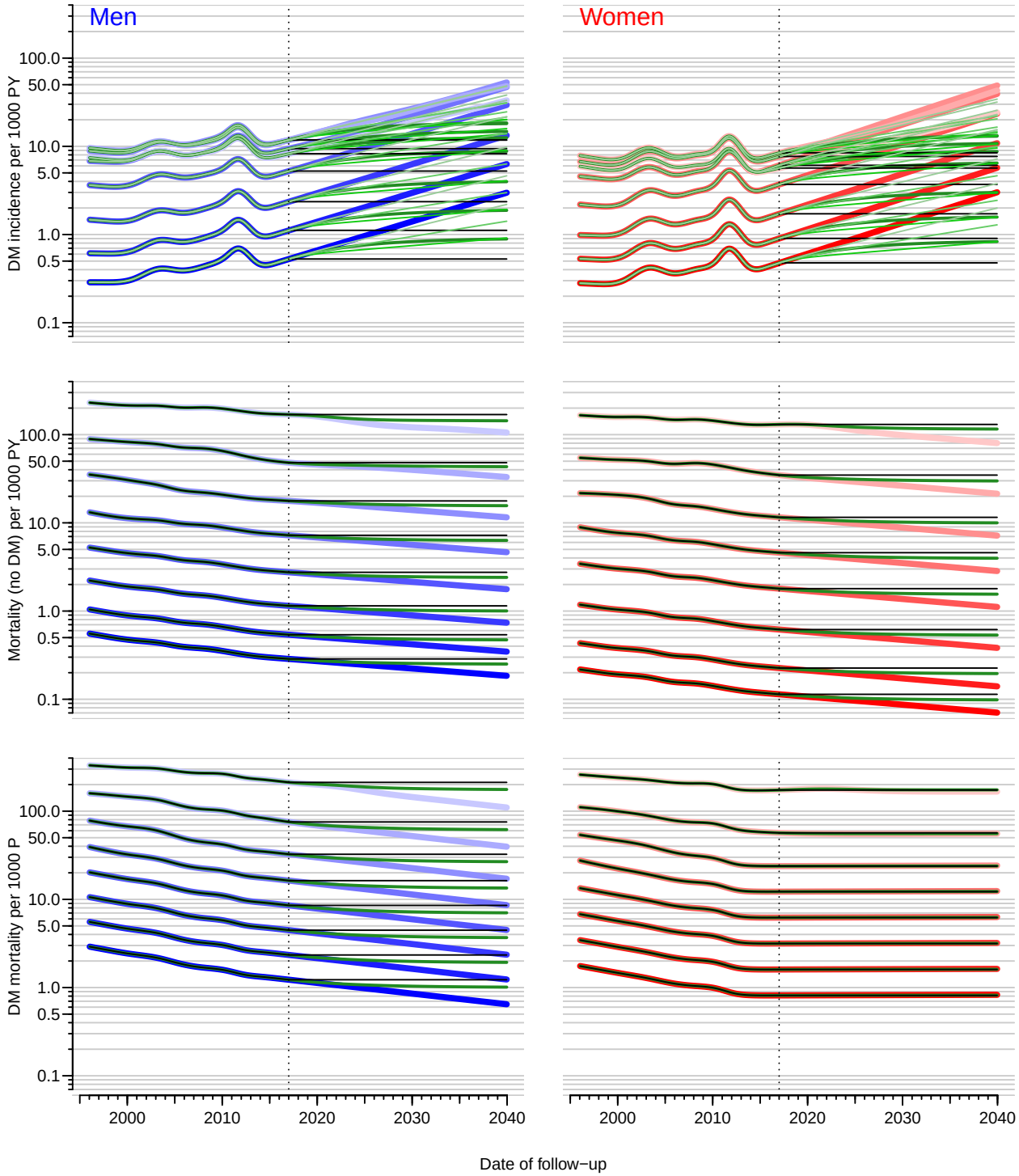


Figure 9.11: *Trends 1996–2016 and projections 2017–2040 of incidence and mortality rates for ages 20 (darkest), 30, . . . , 90 (lightest), as estimated from the age-period-cohort models and projcted by naive extrapolation of the natural splines.*
The thin black lines indicate predictions fixing all rates at the level at the end of 2016, the thick green lines using an attenuation of the trend halving the slope every 4 years, and the thin green lines for incidence rates represent an annual increase of 2.0, 4.0 and 6.0 % per year (dark to light) added (phased in over 1 year).
The vertical dotted lines indicate the end of available data. `./graph/pr-rates-projrates40`

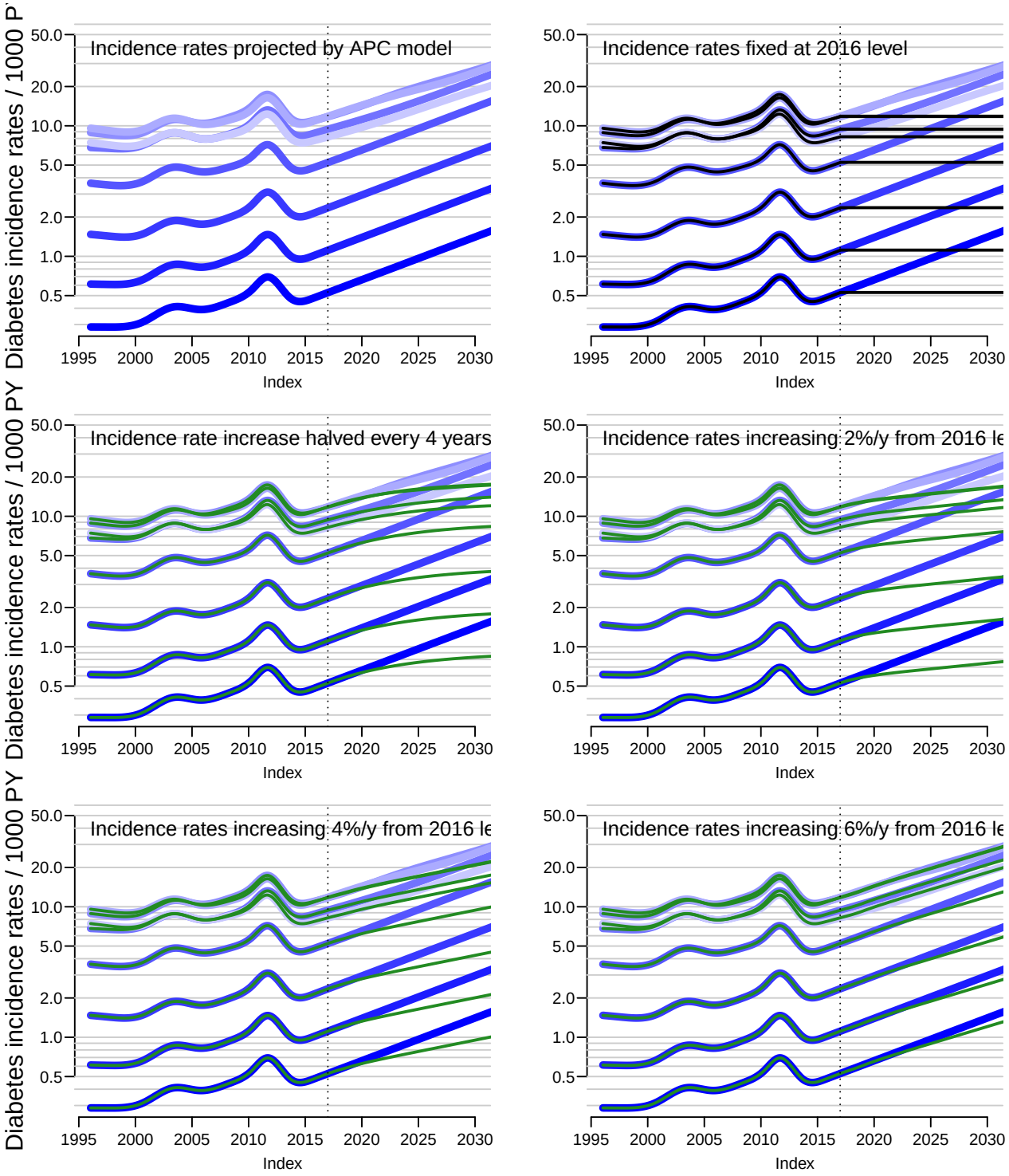


Figure 9.12: *Predicted incidence rates for men, under differerent scenarios*
`./graph/pr-rates-pr-men`

```
> par( mfrow=c(3,2), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> rpl1( Lambda, "W", inc=NULL, yl="Diabetes incidence rates / 1000 PY" )
> text( 1996, 35, "Incidence rates projected by APC model", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "W", inc="fix", icol="black" )
> text( 1996, 35, "Incidence rates fixed at 2016 level", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "W", inc="att", icol="forestgreen", yl="Diabetes incidence rates / 1000 PY" )
> text( 1996, 35, "Incidence rate increase halved every 4 years", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "W", inc="p20", icol="forestgreen" )
> text( 1996, 35, "Incidence rates increasing 2%/y from 2016 level", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "W", inc="p40", icol="forestgreen", yl="Diabetes incidence rates / 1000 PY" )
> text( 1996, 35, "Incidence rates increasing 4%/y from 2016 level", adj=c(0,0), cex=1.3 )
> rpl1( Lambda, "W", inc="p60", icol="forestgreen" )
> text( 1996, 35, "Incidence rates increasing 6%/y from 2016 level", adj=c(0,0), cex=1.3 )


> elapsed()
----------------------
2019-02-18 at 16:15:00
Time elapsed: 00:04:18
----------------------
```

. . . now input from `pred.tex`

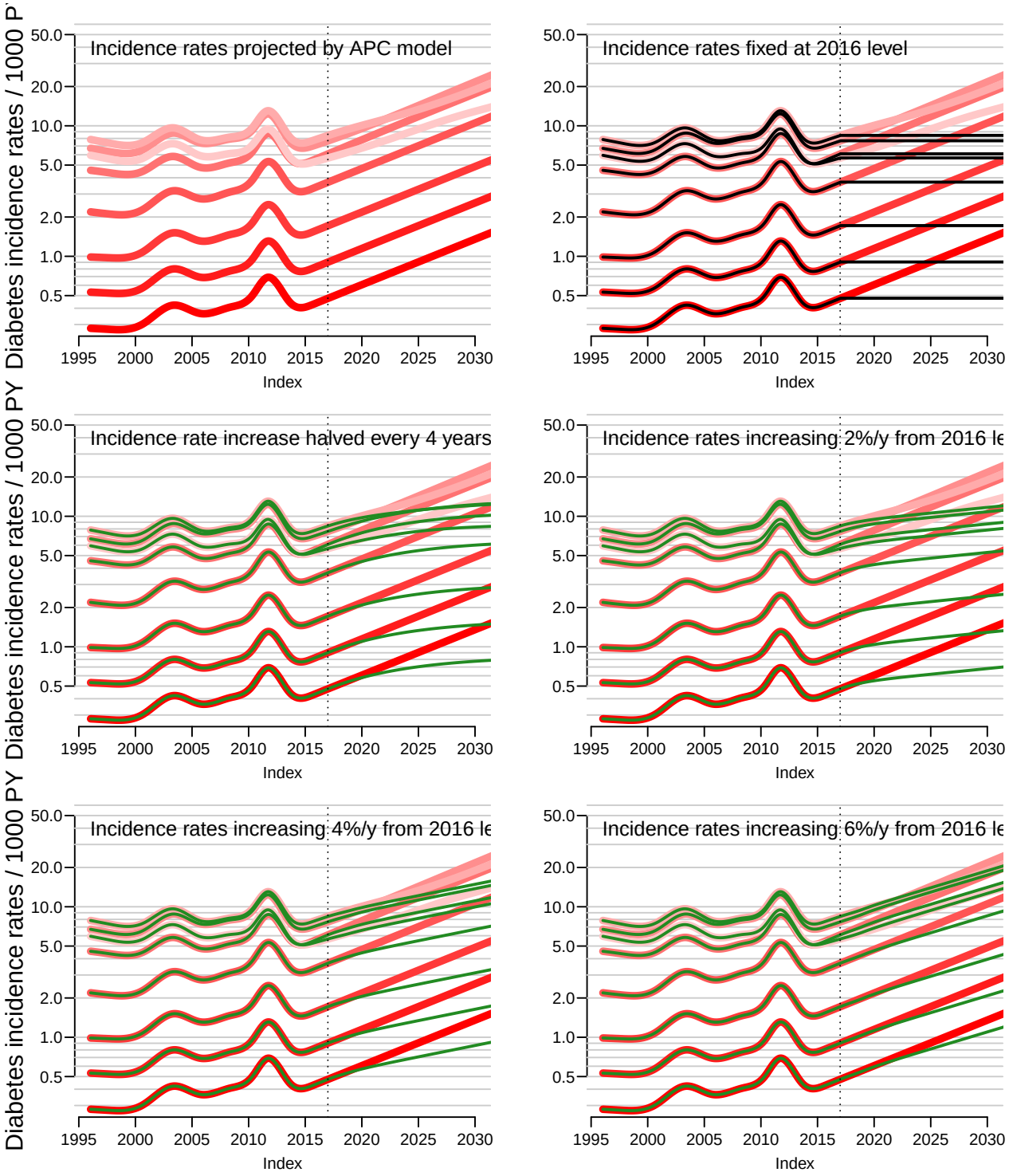


Figure 9.13: *Predicted incidence rates for women, under differerent scenarios*
`./graph/pr-rates-pr-women`

Chapter 10

Predicting prevalence of diabetes

```
> library( Epi )
> start()
---------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-01-11 18:09:58
---------------------------------------
```

In order to predict the prevalent *number* of DM patients in the future, we use the predicted incidence and mortality rates for the period 2016–2040 to predict the prevalence (fraction with DM) and subsequent multiply this with the population predictions from Statistics Denmark.

We use different scenarios for the `incidence` rates:

- Use the naively predicted rates from the APC-model with natural splines — the "`apc`" component of the rate-arrays.

  This will give a prediction of numbers which are presumably the least credible, because of the dramatic decrease seen after 2012 and increase after 2015.

- Use the attenuated rates — the "`att`" component of the rate-arrays.

- Use the rates fixed at the 2017-01-01 level — the "`fix`" component of the rate-arrays.

- Use the rates from 2017-01-01 with an annual increase of 2, 4 and 6% respectively — the "`p20`", "`p40`", "`p60`" components of the rate-arrays.

Furthermore, the whole exercise will be repeated with mortality (both non-DM and DM) declining at the predicted pace and with mortality rates attenuated after 1 January 2016 by a half-time of the slope of 4 years and finally with mortality rates fixed at the 2016 level, so in total $6 \times 3 = 18$ different scenarios.

10.1 Predicted rates

We will start with the observed (smoothed) age-specific prevalences at 2017-01-01 and then use the different scenarios laid out above to predict the prevalences each year till 2040.

First we load the estimated / predicted rates

```
> load( file="../nydata/inits.Rda", v=T )
```

```
Loading objects:
  qn
  fC
  fCp
  fCtable
  cstr
  int
  a.pt
  t.pt
  p.pt
  d.pt
  nk.a
  nk.c
  nk.p
  nk.d
> load( file="../nydata/allrates.Rda", v=T )
Loading objects:
  Lambda
  Mu.DM
  Mu.nD
```

For comparison we also need the number of cases by sex age and calendar time:

```
> load( file="../nydata/prevN.Rda", v=T )
Loading objects:
  prN
> str( prN )
'data.frame':        22000 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","W": 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 2 3 2 0 3 6 1 3 ...
 $ T2 : num  0 0 0 0 0 0 0 0 0 0 ...
 $ nD : num  3822 3939 3930 3872 3685 ...
 $ N  : num  3822 3939 3932 3875 3687 ...
 $ DM : num  0 0 2 3 2 0 3 6 1 3 ...
```

We shall use a multistate calculation scheme to predict the course of DM prevalence development in the future population under various scenarios of mortality and incidence development. So we set up structures to help calculations and hold results:

`pr.fit` — array of predicted age-specific prevalences at 1995-01-01 to 2017-01-01, smoothed by natural splines. This was derived in the section on prevalence, however not as finely as we need, so we load the relevant models in order to be able to make the relevant predictions:

```
> load( file="../nydata/prevalences.Rda", v=T )
Loading objects:
  parr
  mod
  akn
  pr.obs
  pr.ini
  pr.fin
```

```
> str( parr )

 num [1:2, 1:3, 1:3, 1:201, 1:22, 1:3] 3.42e-04 1.26e-04 5.78e-06 6.62e-06 5.35e-04 ...
 - attr(*, "dimnames")=List of 6
  ..$ mod: chr [1:2] "glm" "gam"
  ..$ typ: chr [1:3] "T1" "T2" "DM"
  ..$ sex: chr [1:3] "M" "W" "B"
  ..$ A  : chr [1:201] "0" "0.5" "1" "1.5" ...
  ..$ T  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$    : chr [1:3] "Est" "lo" "hi"

> pr.fit <- parr["glm","DM",,1:100*2,,"Est"]
> str( pr.fit )

 num [1:3, 1:100, 1:22] 0.000559 0.000462 0.00051 0.00061 0.000507 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex: chr [1:3] "M" "W" "B"
  ..$ A  : chr [1:100] "0.5" "1.5" "2.5" "3.5" ...
  ..$ T  : chr [1:22] "1996" "1997" "1998" "1999" ...

> dimnames( pr.fit )[[3]]

 [1] "1996" "1997" "1998" "1999" "2000" "2001" "2002" "2003" "2004" "2005" "2006" "2007"
[13] "2008" "2009" "2010" "2011" "2012" "2013" "2014" "2015" "2016" "2017"
```

`TR` — array of transition probabilities between states no-DM and DM (omitting Death). Transition probabilities are computed under the 18 different extrapolation scenarios. The rates that form the base for these refer to (midpoints of) date intervals of length `int` and are therefore labeled on the period dimension by the midpoint of these, a total of (2040-2017-1)/`int`=264. The labeling along the age-dimension is by the border between the age-intervals

`prv` — array of predicted prevalences based on the initial prevalences at 2017-01-01 and the transition probabilities as put in `TR`. Note that we use 2 scenario dimensions, one referring to the 6 scenarios for *incidence* rate prediction:

- `apc` — naive predictions of rates by extrapolating the linear part of the natural splines.
- `att` — attenuation of the slopes predicted, halving the slope of the predictions every 3 years.
- `fix` — fixing rates at the level of 1 January 2017.
- `p20` — sharp attenuation of the slopes of predicted rates with a slope half time of 6 months, and adding an extra increase of 2% per year, increasing from 0 to full 2% after 2 years.
- `p40` — same as above, but with 4% added.
- `p60` — same as above, but with 6% added.

. . . and the other dimension referring to the three projection scenarios for *mortality* rates (using the same scenarios for mortality rates among persons with and without diabetes):

- `apc` — continuing the trend from the spline models, essentially amounting to a continuing decline in mortality.

`att` — attenuating the decline with a half-time of 4 years; that is the *slope* in mortality is halved every 4 years — over the 13 years prediction period 1917–2030 this is a reduction to $0.5^{13/4} = 0.105$ of the original slope, and over the 23 years prediction period t0 2040 this is a reduction to $0.5^{23/4} = 0.019$ of the original slope.

`fix` — fixing rates at the level of 1 January 2017.

`prn` — array of predicted *number* of DM patients in one-year age classes at the 1 January each year. So the same structure as `prv`, but with substantially fewer entries — namely 1-year age-classes and 1 year spaced dates. Obtained by multiplying age-averages at 1 January each year by the population figures obtained from Statistics Denmark.

10.2 Transition probabilities

In order to get the predicted prevalences (fractions) of DM by age, period and prediction type, we need the (1-step) transition matrices at all combinations of age (a) and date (p), this is put in an array. But we will only need the rates from 2017 and onward, so we restrict the arrays with the rates to this period, and also to the relevant scenarios:

```
> dimnames( Lambda )[[2]][252+0:1]
[1] "2016.95833333333" "2017.04166666667"
> dimnames( Lambda )[[4]]
[1] "ap"  "apc" "gam" "LCa" "att" "fix" "p20" "p40" "p60"
> dimnames( Mu.nD  )[[4]]
[1] "ap"  "apc" "gam" "LCa" "att" "fix" "p20" "p40" "p60"
> rLambda <- Lambda[,-(1:252),,c("apc","att","fix","p20","p40","p60")]
> rMu.nD  <-  Mu.nD[,-(1:252),,c("apc","att","fix")]
> rMu.DM  <-  Mu.DM[,-(1:252),,c("apc","att","fix")]
> states <- c("nD","DM")
> TR <- NArray( c( dimnames( rLambda )[-4],
+                  list( imod = dimnames( rLambda )[[4]],
+                        mmod = dimnames( rMu.nD  )[[4]],
+                        from = states,
+                          to = states ) ) )
> str( TR ) ; fCp( length( TR ) )
 logi [1:1200, 1:276, 1:2, 1:6, 1:3, 1:2, 1:2] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 7
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ p   : chr [1:276] "2017.04166666667" "2017.125" "2017.20833333333" "2017.29166666667" ..
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
  ..$ from: chr [1:2] "nD" "DM"
  ..$ to  : chr [1:2] "nD" "DM"
[1] 47,692,800
> str( rLambda ) ; fCp( length( rLambda ) )
 num [1:1200, 1:276, 1:2, 1:6] 9.94e-06 1.00e-05 1.01e-05 1.01e-05 1.02e-05 ...
 - attr(*, "dimnames")=List of 4
  ..$ a  : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" ..
  ..$ p  : chr [1:276] "2017.04166666667" "2017.125" "2017.20833333333" "2017.29166666667" ...
  ..$ sex: chr [1:2] "M" "W"
  ..$ mod: chr [1:6] "apc" "att" "fix" "p20" ...
```

```
[1] 3,974,400
```

In order to fill `TR`, we need the cumulative incidences over intervals of length `int`. But these were exactly what we predicted in the previous sections by setting the person years equal to `int`. Note, both age and period are in 1 month intervals:

```
> summary( diff(as.numeric(dimnames(rLambda)[["a"]]) ) )
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.08333 0.08333 0.08333 0.08333 0.08333 0.08333
> summary( diff(as.numeric(dimnames(rLambda)[["p"]]) ) )
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.08333 0.08333 0.08333 0.08333 0.08333 0.08333
> int
[1] 0.08333333
```

So we can now compute the one-step transition matrices for every combination of `a.pt` and `p.pt`. Note that we are using the small-interval approximation to the survival: $\exp(-\Lambda - M_{\mathrm{nD}}) \approx 1 - \Lambda - M_{\mathrm{nD}}$ and Λ as the transition probability from nD to DM. This only works because we operate with small intervals where the approximation is good, and in particular where the probability of two transitions in one intervals is negligible

```
> for( ii in dimnames(TR)[["imod"]] )
+ for( im in dimnames(TR)[["mmod"]] )
+    {
+ TR[,,,ii,im,"nD","nD"] <- 1 - (rLambda[,,,ii]+rMu.nD[,,,im])
+ TR[,,,ii,im,"nD","DM"] <-      rLambda[,,,ii]
+ TR[,,,ii,im,"DM","nD"] <- 0
+ TR[,,,ii,im,"DM","DM"] <- 1 -                 rMu.DM[,,,im]
+    }
```

Note that we have not included the "Dead" state in the calculations, because we only bother about the *fraction* of diabetes patients in each age class at each time-point. So the probabilities we compute do not sum to 1 within the "from" states; we only compute how many of the persons alive that end up being alive at the next time point — and in what state they are at that point.

We have now collected the transition probabilities between "nD" and "DM" as well as the probabilities of remaining in each of these two states, all referring to a duration of `int`.

```
> save( TR, file="../nydata/TRf.Rda" )
> load(     file="../nydata/TRf.Rda" )
```

10.3 Prediction of the observed prevalences

We do not need to predict the population size; we can get away with only predicting the prevalences as fractions. When we multiply the fraction of persons in states (nD,DM) with the transition matrix, we get the fraction of the persons alive at the beginning of the interval (as either nD or DM) that are in states (nD,DM) at the end of the interval. These do not sum to 1 (because the ones dying in the interval are left out), so we rescale to prevalence in each step.

First we set up an array to hold the predicted prevalences under different scenarios:

```
> dpr <- dimnames(TR)[1:5]
> names( dpr )[2] <- "t"
> dpr[[2]] <- t.pt[t.pt>=2017]
> prv <- NArray( dpr )
> str( prv ) ; fCp( length( prv ) )
 logi [1:1200, 1:277, 1:2, 1:6, 1:3] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ t   : chr [1:277] "2017" "2017.08333333333" "2017.16666666667" "2017.25" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
[1] 11,966,400
```

Thus we must make a loop that updates the prevalences at 2017-01-01 to those at subsequent dates, (that is next time, next age) but first we must initialize the prevalences as modeled on 2017-01-01, as well as the prevalences at ages 0 (the new-born) that we set to 0; note that we exploit the column major storage of arrays to get `prv` filled with identical values across the two last dimensions (incidence and mortality scenarios).

```
> # Smoothed prevalences at 2017-01-01 - the starting values
> # Repeated by virtue of the column major storage of arrays
> for( sx in c("M","W") ) prv[,"2017",sx,,] <- pr.fin["DM",sx,,"Est"]
> # Prevalences at age 0:
> prv[1,,,,] <- 0
```

From these initial values we can compute the predicted prevalences under the different scenarios. We take the fraction of the population in age class `ia` at time `ip` that end up as diabetes patients at time `ip+1` (and hence in age class `ia+1`), and divide by the fraction of all that remain alive, which is the diabetes patients, *plus* those who survive free of diabetes:

```
> str( prv )
 num [1:1200, 1:277, 1:2, 1:6, 1:3] 0 0.00109 0.0011 0.0011 0.00111 ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ t   : chr [1:277] "2017" "2017.08333333333" "2017.16666666667" "2017.25" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
> str( TR )
 num [1:1200, 1:276, 1:2, 1:6, 1:3, 1:2, 1:2] 1 1 1 1 1 ...
 - attr(*, "dimnames")=List of 7
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ p   : chr [1:276] "2017.04166666667" "2017.125" "2017.20833333333" "2017.29166666667" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
  ..$ from: chr [1:2] "nD" "DM"
  ..$ to  : chr [1:2] "nD" "DM"
> system.time(
+ for( ip in 1:(dim(prv)[2]-1) )
+ for( ia in 1:(dim(prv)[1]-1) )
+ prv[ia+1,ip+1,,,] <- (    prv[ia,ip,,,]  * TR[ia,ip,,,,"DM","DM"]
+                       +(1-prv[ia,ip,,,]) * TR[ia,ip,,,,"nD","DM"] ) /
```

```
+                      (    prv[ia,ip,,,]  * TR[ia,ip,,,,"DM","DM"]
+                       +(1-prv[ia,ip,,,]) * TR[ia,ip,,,,"nD","DM"]
+                       +(1-prv[ia,ip,,,]) * TR[ia,ip,,,,"nD","nD"] )
+             )
   user  system elapsed
  16.27    0.04   16.31
```

Note that the code above is particularly simple because we only need to compute the prevalence at the next date and age. If we had had a more elaborate model with, say complications states, the calculations in the loop would have been a matrix-multiplication updating the state-distribution, but this simplification would have been at the expense of another three loop-levels, namely over the three last dimensions of the `prv` array.

We can then show a few of the predicted age-specific prevalences

```
> pplt <-
+ function( sex, per, isc, msc )
+    {
+ np <- length(per)
+ matplot( a.pt, prv[,per,sex,isc,msc]*100,
+          type="l", lty=1, lwd=2, col=gray((1:np+1)/(np+5)),
+          xlim=c(20,90), ylim=c(0,45), xaxt="n", yaxt="n", yaxs="i" )
+ abline( h=seq(0,50,5), v=seq(0,90,10), col=gray(0.7), lty="11", lend="butt" )
+ text( 22, 44, paste( sx, "\nInc:", isc, "\nMort:", msc ), adj=c(0,1) )
+    }
> par( mfcol=c(5,4), mar=c(1,1,0,0), oma=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> for( sx in c("M","W") )
+ for( ms in c("apc","fix") )
+    {
+ for( is in c("att","fix","p20","p40","p60") )
+    {
+    pplt( sx, paste(seq(2017,2038,3)), is, ms )
+    if(sx=="M" & ms=="apc") axis(side=2)
+    }
+    axis( side=1 )
+    }
> mtext( "Age (years)", side=1, outer=TRUE, line=1 )
> mtext( "Prevalence (%)", side=2, outer=TRUE, line=1, las=0 )

> save( prv, file="../nydata/prv-pred.Rda" )
> load(      file="../nydata/prv-pred.Rda" )
```

10.4 The actual numbers of diabetes patients in Denmark

In the previous section we only looked at the age-specific prevalences, because these are the quantities that are driven by the incidence and mortality rates. However, it is also of interest to see how the *number* of diabetes patients would have looked under the different scenarios.

To show the number of patients we set up an array `prn` with *structure* (but not extent) as `prv` to hold the *number* of diabetes patients by category, assuming the age-distribution in the population to be as actually observed (that is as extracted from Statistics Denmark). However `prn` will have 100 age-classes rather than 1200 (100/`int`), and only 23 dates (2018–2040): `prv`. This is because we have the predicted population size in 1-year classes.

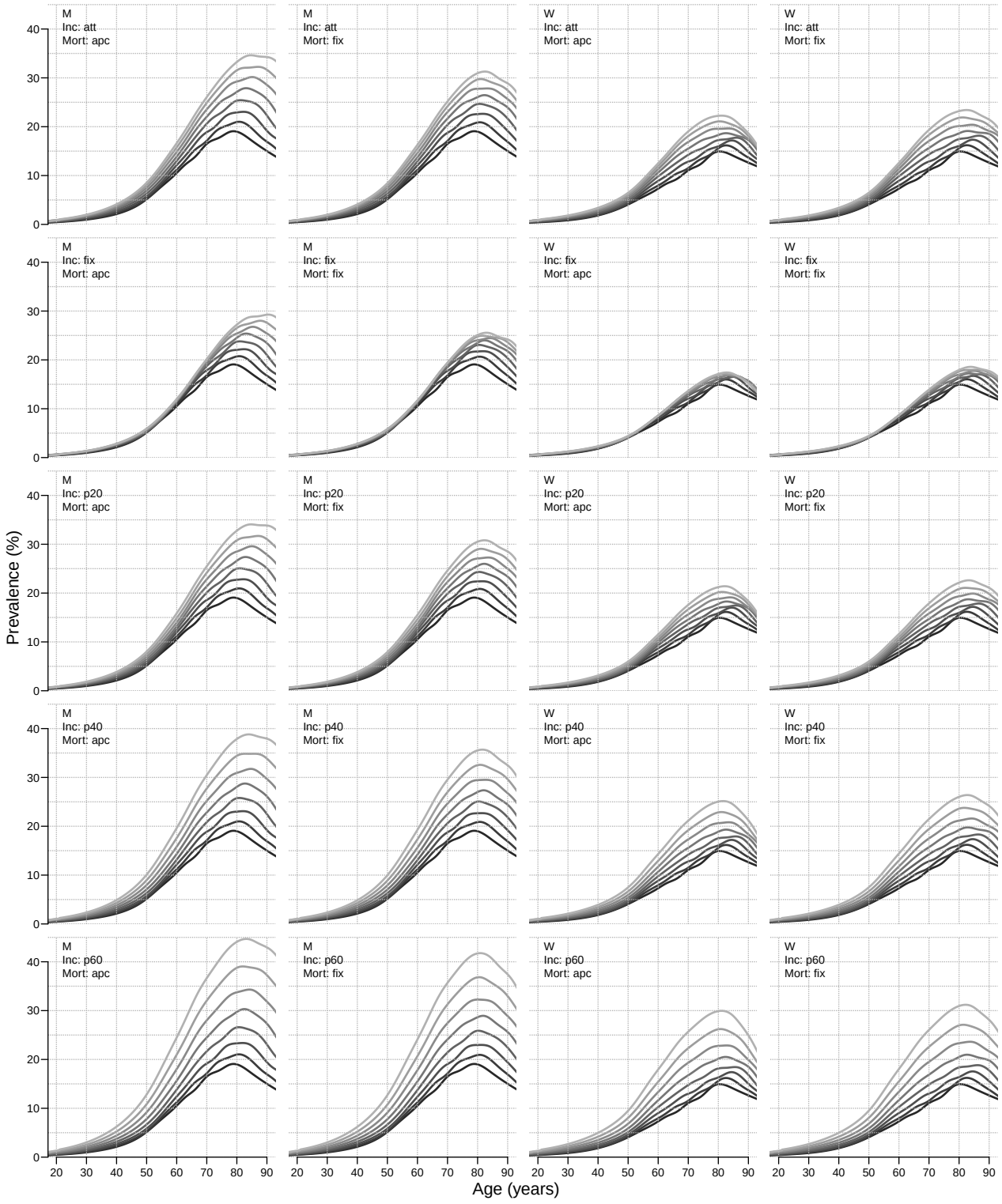


Figure 10.1: *Predicted age-specific prevalences 2017–38 by 3 years for men and women under different scenarios. Colouring is from dark (2017) to light (2038).* `./graph/pred-a-prv`

```
> dn <- dimnames(prv)
> dn[[1]] <- 0:99
> dn[[2]] <- 2018:2040
> prn <- NArray( dn )
> str( prv ) ; fCp( length(prv) )
 num [1:1200, 1:277, 1:2, 1:6, 1:3] 0 0.00109 0.0011 0.0011 0.00111 ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ t   : chr [1:277] "2017" "2017.08333333333" "2017.16666666667" "2017.25" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"

[1] 11,966,400

> str( prn ) ; fCp( length(prn) )

 logi [1:100, 1:23, 1:2, 1:6, 1:3] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:100] "0" "1" "2" "3" ...
  ..$ t   : chr [1:23] "2018" "2019" "2020" "2021" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"

[1]    82,800
```

In order to fill in the numbers we use the estimated age-specific prevalences at 1st January each year, that is at the dates 2018-01-01,...,2040-01-01 in the entries along the `t`-dimension of `prv`. Moreover we want the prevalences for a 1 year age class rather than age-classes of length `int`. So we take the average prevalences from `prv` over each one-year age-interval.

This is really just a simple matrix operation; take a diagonal matrix of 1/12 (well, `int`), and repeat each column 12 (1/`int`) times:

```
> dd <- diag(100)[,rep(1:100,each=1/int)]*int
> print( round( as.table(dd[1:3,1:16]), 3 ), zero.print="." )
      A     B     C     D     E     F     G     H     I     J     K     L     M     N
A 0.083 0.083 0.083 0.083 0.083 0.083 0.083 0.083 0.083 0.083 0.083 0.083     .     .
B     .     .     .     .     .     .     .     .     .     .     .     . 0.083 0.083
C     .     .     .     .     .     .     .     .     .     .     .     .     .     .
      O     P
A     .     .
B 0.083 0.083
C     .     .
```

Pre-multiplying this 100×1200 matrix to the 1200 ($= 100/$`int`) age-specific prevalences gives the average prevalences in the 100 1-year age-classes. So we just select the dates at which we want the prevalences:

```
> dimnames(prv)[[2]][wh<-seq(13,277,12)]
 [1] "2018" "2019" "2020" "2021" "2022" "2023" "2024" "2025" "2026" "2027" "2028" "2029"
[13] "2030" "2031" "2032" "2033" "2034" "2035" "2036" "2037" "2038" "2039" "2040"

> for( sx in dimnames(prn)[[3]] )
+ for( im in dimnames(prn)[[4]] )
+ for( mm in dimnames(prn)[[5]] )
+ prn[,,sx,im,mm] <- dd %*% prv[,wh,sx,im,mm]
```

Now `prn` contains the prevalences (as fractions $\in [0,1]$) for 100 age classes and the 24 dates 2017-01-01 through 2040-01-01 (for each combination of sex and prediction assumptions for incidences and mortalities, respectively). We need to multiply these prevalences by the population figures for each age, date and sex. This is in the array `beff`:

```
> load( "../nydata/pop.Rda" )
> str( pop <- xtabs( N ~ A + P + sex, data=beff ) )
 'xtabs' int [1:100, 1:24, 1:2] 30054 28889 28451 28302 29187 30169 32201 31897 32974 32724 ..
 - attr(*, "dimnames")=List of 3
  ..$ A  : chr [1:100] "0" "1" "2" "3" ...
  ..$ P  : chr [1:24] "2017" "2018" "2019" "2020" ...
  ..$ sex: chr [1:2] "W" "M"
 - attr(*, "call")= language xtabs(formula = N ~ A + P + sex, data = beff)
> str( prn )
 num [1:100, 1:23, 1:2, 1:6, 1:3] 5.86e-05 1.16e-03 1.35e-03 1.45e-03 1.56e-03 ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:100] "0" "1" "2" "3" ...
  ..$ t   : chr [1:23] "2018" "2019" "2020" "2021" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
> round( range( prn ), 5 )
[1] 0.00006 0.48779
```

So now `pop` is an array with dimensions equal to the 3 first dimensions of `prn`, hence we loop over the last three dimensions when multiplying with the population size:

```
> dmp <- prn * 0
> for( sx in dimnames(dmp)[[3]] )
+ for( ii in dimnames(dmp)[[4]] )
+ for( im in dimnames(dmp)[[5]] )
+      dmp[,,sx,ii,im] <- prn[,,sx,ii,im] * pop[,dimnames(prn)[[2]],sx]
```

The array `dmp` now holds the predicted *number* of diabetes patients in Denmark by sex, age (1-year classes) and date (1-year equidistance), for 18 different prediction scenarios.

First we draw simple population pyramids of the age-distribution of the diabetes patients in Denmark, as predicted under different scenarios. Note that this uses the undocumented feature of `barplot` that if the first number in a column is negative this is taken as the left endpoint of the bar. So `c(-m,m,f)` is a bar starting at `-m`, and a division at `-m+m`(=0) and an upper end at `-m+m-f`. Coloring is from the top, that is the part stretching from `-m+m` to `-m+m+f` get the first color

```
> clr <- c("red","blue")
> draw.dmp <-
+ function(pp,wh.i,wh.m,lim=6)
+ {
+ par( mar=c(3,3,3,0), mgp=c(3,1,0)/1.6, las=1 )
+ barplot( height=t( cbind( -dmp[,pp,"M",wh.i,wh.m],
+                            dmp[,pp,"M",wh.i,wh.m],
+                            dmp[,pp,"W",wh.i,wh.m] ) )/ 1000,
+          horiz=TRUE, col=clr,
+          border="transparent",space=0,axes=FALSE,
+          names.arg=rep("",dim(prn)[1]),
+          xlim=c(-1,1)*lim*1.05,
```

```
+          xlab="Persons in 1 year class (1000s)",ylab="Age")
+ abline(h=seq(0,100,5),
+        v=seq(-lim,lim,0.5),
+        col="white")
+ text( 6, 20, paste("inc:", wh.i, "\nmort:", wh.m), adj=c(1,1) )
+ axis( side=1, at=seq(-lim,lim,1), labels=abs(seq(-lim,lim,1)) )
+ axis( side=1, at=seq(-lim,lim,0.5), labels=NA, tcl=-0.3 )
+ axis( side=2, at=seq(0,100,20) )
+ axis( side=2, at=seq(0,100,5), labels=NA, tcl=-0.3 )
+ mtext( pp, at=-lim, adj=1.2, cex=1.0, font=2 )
+ mtext( formatC(sum(dmp[,pp,"M",wh.i,wh.m]),0,format="f",big.mark=","),
+        at=-1, col="blue", line=0, cex=0.99, adj=1 )
+ mtext( formatC(sum(dmp[,pp,"W",wh.i,wh.m]),0,format="f",big.mark=","),
+        at= 1, col="red" , line=0, cex=0.99, adj=0 )
+ mtext( "N", at=0, line=0, cex=0.99 )
+ }
> pdf( "pred-fix-film.pdf", width=8, height=6 )
> for( pp in paste(2018:2040) ) draw.dmp(pp,"fix","apc",lim=10)
> dev.off()
null device
          1


> pdf( "pred-att-film.pdf", width=8, height=6 )
> for( pp in paste(2018:2040) ) draw.dmp(pp,"att","att",lim=10)
> dev.off()
null device
          1


> par( mfrow=c(2,2), mar=c(3,3,0,0), oma=c(0,0,2,0), mgp=c(3,1,0)/1.6, las=1 )
> draw.dmp("2018","att","att",lim=8)
> draw.dmp("2022","att","att",lim=8)
> draw.dmp("2026","att","att",lim=8)
> draw.dmp("2030","att","att",lim=8)
> mtext( "Incidence rate increase and mortality decrease attenuated",
+        side=3, line=0, outer=TRUE )
```

10.5 Time trends in prevalent number of DM patients

First we make a table of the total predicted number of DM patients by date, sex and scenario:

```
> str( dmp )
 num [1:100, 1:23, 1:2, 1:6, 1:3] 1.9 37.5 41.3 43.9 46.6 ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:100] "0" "1" "2" "3" ...
  ..$ t   : chr [1:23] "2018" "2019" "2020" "2021" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
> dimnames( dmp )[[4]]
[1] "apc" "att" "fix" "p20" "p40" "p60"
```

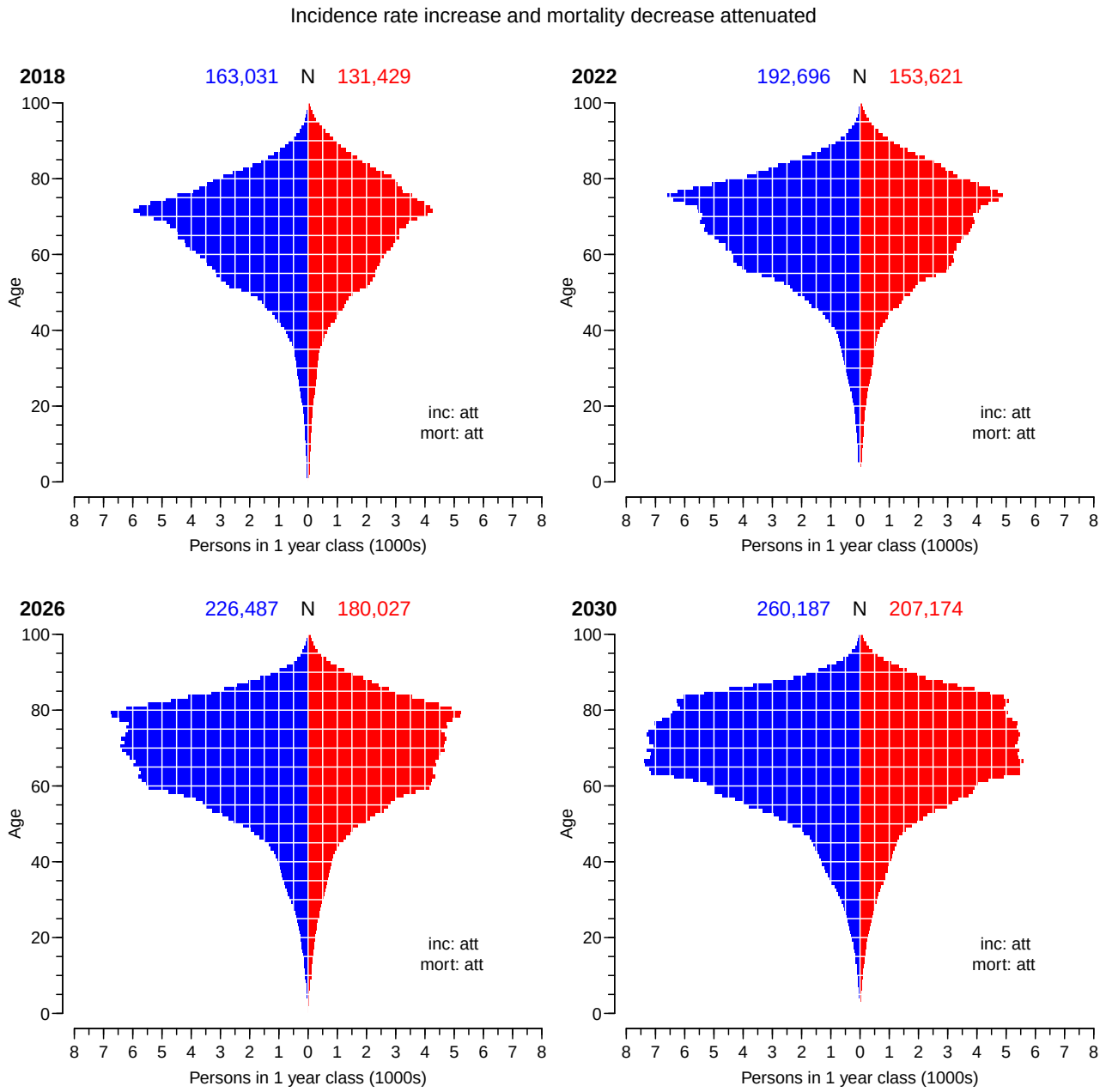


Figure 10.2: *Predicted number and sex- and age-distribution of prevalent diabetes patients in Denmark, 2018–2030. The prediction scenario used for incidence and mortality rates is a linear prediction of log-rates from age-period-cohort models with natural splines, where the estimated change by calendar time is halved every 4 years.* `./graph/pred-att-att`

```
> fCtable( addmargins( round( apply( dmp, 2:5, sum ) ), 2 ),
+          col.vars=3, row.vars=c(2:1,4), d=0, w=10 )
```

		imod	apc	att	fix	p20	p40	p60
sex	t	mmod						
M	2018	apc	163,048	163,033	162,697	162,998	163,015	163,033
		att	163,046	163,031	162,695	162,996	163,014	163,031
		fix	163,011	162,996	162,660	162,961	162,978	162,996
	2019	apc	169,934	169,800	168,439	169,570	169,726	169,884
		att	169,921	169,787	168,426	169,557	169,713	169,871

	fix	169,780	169,646	168,285	169,416	169,572	169,730
2020	apc	177,547	177,081	174,072	176,465	177,000	177,548
	att	177,504	177,038	174,029	176,421	176,956	177,504
	fix	177,194	176,728	173,721	176,112	176,647	177,195
2021	apc	185,924	184,800	179,591	183,589	184,781	186,020
	att	185,823	184,699	179,491	183,488	184,680	185,918
	fix	185,290	184,167	178,962	182,956	184,148	185,386
2022	apc	195,111	192,890	184,996	190,886	193,024	195,276
	att	194,917	192,696	184,806	190,694	192,831	195,082
	fix	194,114	191,894	184,012	189,893	192,028	194,278
2023	apc	205,142	201,273	190,276	198,313	201,694	205,303
	att	204,815	200,947	189,956	197,989	201,368	204,975
	fix	203,696	199,832	188,858	196,877	200,253	203,856
2024	apc	216,056	209,878	195,422	205,844	210,773	216,104
	att	215,546	209,372	194,928	205,341	210,266	215,594
	fix	214,076	207,908	193,493	203,883	208,802	214,122
2025	apc	227,899	218,646	200,432	213,466	220,257	227,698
	att	227,155	217,909	199,718	212,735	219,519	226,953
	fix	225,298	216,064	197,918	210,901	217,672	225,094
2026	apc	240,710	227,509	205,296	221,162	230,135	240,094
	att	239,674	226,487	204,310	220,148	229,110	239,056
	fix	237,401	224,234	202,124	217,913	226,854	236,781
2027	apc	254,536	236,417	210,010	228,923	240,405	253,313
	att	253,145	235,050	208,700	227,570	239,033	251,921
	fix	250,431	232,368	206,112	224,913	236,346	249,203
2028	apc	269,426	245,321	214,571	236,743	251,070	267,377
	att	267,616	243,549	212,885	234,993	249,290	265,565
	fix	264,441	240,422	209,884	231,899	246,153	262,385
2029	apc	285,430	254,179	218,980	244,617	262,128	282,305
	att	283,134	251,942	216,863	242,410	259,876	280,006
	fix	279,481	248,357	213,444	238,869	256,277	276,349
2030	apc	302,596	262,951	223,232	252,537	273,578	298,114
	att	299,745	260,187	220,633	249,815	270,791	295,261
	fix	295,602	256,140	216,795	245,822	266,722	291,116
2031	apc	320,969	271,602	227,325	260,496	285,415	314,818
	att	317,494	268,251	224,193	257,201	282,031	311,344
	fix	312,855	263,739	219,941	252,755	277,488	306,705
2032	apc	340,600	280,105	231,262	268,492	297,644	332,435
	att	336,436	276,112	227,552	264,570	293,602	328,276
	fix	331,300	271,137	222,892	259,676	288,586	323,142
2033	apc	361,535	288,433	235,041	276,519	310,258	350,975
	att	356,618	283,743	230,709	271,917	305,502	346,069
	fix	350,985	278,313	225,655	266,582	300,017	340,443
2034	apc	383,832	296,576	238,671	284,583	323,270	370,463
	att	378,106	291,142	233,681	279,255	317,746	364,754
	fix	371,980	285,264	228,246	273,488	311,798	358,642
2035	apc	407,544	304,518	242,158	292,685	336,682	390,913
	att	400,956	298,297	236,477	286,589	330,343	384,353
	fix	394,343	291,981	230,673	280,399	323,940	377,762
2036	apc	432,737	312,264	245,516	300,839	350,515	412,357
	att	425,242	305,217	239,116	293,938	343,318	404,903
	fix	418,149	298,471	232,957	287,333	336,464	397,840
2037	apc	459,457	319,803	248,747	309,045	364,769	434,805
	att	451,018	311,897	241,606	301,306	356,676	426,421
	fix	443,452	304,730	235,104	294,293	349,378	418,897
2038	apc	487,772	327,147	251,871	317,322	379,468	458,287
	att	478,357	318,351	243,968	308,713	370,445	448,947

		fix	470,330	310,771	237,135	301,302	362,710	440,972
	2039	apc	517,728	334,298	254,896	325,677	394,624	482,818
		att	507,318	324,586	246,215	316,172	384,641	472,502
		fix	498,844	316,602	239,062	308,369	376,475	464,090
	2040	apc	549,372	341,264	257,834	334,121	410,248	508,410
		att	537,954	330,611	248,358	323,695	399,279	497,106
		fix	529,050	322,232	240,898	315,508	390,691	488,275
W	2018	apc	131,442	131,428	131,137	131,396	131,409	131,423
		att	131,442	131,429	131,138	131,397	131,410	131,423
		fix	131,460	131,447	131,156	131,415	131,428	131,442
	2019	apc	136,486	136,368	135,181	136,149	136,269	136,389
		att	136,492	136,375	135,187	136,156	136,275	136,396
		fix	136,564	136,446	135,258	136,227	136,346	136,467
	2020	apc	142,154	141,741	139,105	141,138	141,549	141,970
		att	142,177	141,763	139,126	141,160	141,571	141,992
		fix	142,330	141,917	139,279	141,313	141,724	142,145
	2021	apc	148,483	147,478	142,893	146,271	147,188	148,141
		att	148,534	147,529	142,943	146,321	147,239	148,192
		fix	148,795	147,789	143,201	146,581	147,498	148,452
	2022	apc	155,528	153,524	146,548	151,498	153,146	154,883
		att	155,625	153,621	146,643	151,594	153,243	154,980
		fix	156,014	154,009	147,026	151,981	153,631	155,369
	2023	apc	163,335	159,816	150,061	156,784	159,396	162,185
		att	163,498	159,978	150,220	156,946	159,558	162,349
		fix	164,036	160,514	150,745	157,479	160,093	162,886
	2024	apc	171,947	166,285	153,417	162,100	165,914	170,043
		att	172,201	166,536	153,661	162,349	166,165	170,297
		fix	172,905	167,235	154,342	163,043	166,863	171,000
	2025	apc	181,416	172,870	156,608	167,426	172,688	178,463
		att	181,787	173,236	156,961	167,788	173,054	178,833
		fix	182,672	174,112	157,809	168,657	173,930	179,718
	2026	apc	191,798	179,519	159,630	172,752	179,714	187,458
		att	192,316	180,027	160,116	173,255	180,224	187,976
		fix	193,398	181,094	161,142	174,309	181,291	189,056
	2027	apc	203,152	186,179	162,478	178,069	186,988	197,045
		att	203,851	186,861	163,123	178,740	187,672	197,742
		fix	205,144	188,128	164,333	179,991	188,942	199,031
	2028	apc	215,539	192,804	165,149	183,367	194,508	207,236
		att	216,455	193,690	165,980	184,237	195,397	208,148
		fix	217,972	195,168	167,379	185,692	196,880	209,659
	2029	apc	229,016	199,348	167,636	188,636	202,265	218,046
		att	230,187	200,471	168,679	189,735	203,396	219,210
		fix	231,940	202,167	170,272	191,401	205,099	220,954
	2030	apc	243,656	205,782	169,948	193,878	210,269	229,500
		att	245,124	207,174	171,229	195,238	211,675	230,955
		fix	247,124	209,095	173,018	197,120	213,608	232,943
	2031	apc	259,511	212,061	172,076	199,079	218,508	241,605
		att	261,320	213,758	173,622	200,732	220,226	243,393
		fix	263,580	215,909	175,607	202,835	222,396	245,634
	2032	apc	276,656	218,166	174,029	204,243	226,992	254,387
		att	278,851	220,202	175,866	206,223	229,060	256,553
		fix	281,382	222,587	178,048	208,550	231,473	259,057
	2033	apc	295,147	224,068	175,807	209,365	235,716	267,861
		att	297,780	226,478	177,960	211,703	238,174	270,451
		fix	300,592	229,101	180,338	214,257	240,836	273,227
	2034	apc	315,059	229,755	177,420	214,449	244,690	282,055
		att	318,184	232,576	179,915	217,182	247,581	285,118

	fix	321,286	235,438	182,486	219,964	250,496	288,175
2035	apc	336,463	235,214	178,875	219,501	253,924	296,995
	att	340,134	238,481	181,736	222,661	257,289	300,584
	fix	343,533	241,581	184,495	225,669	260,458	303,924
2036	apc	359,437	240,444	180,188	224,531	263,433	312,720
	att	363,711	244,192	183,437	228,152	267,316	316,885
	fix	367,411	247,527	186,377	231,383	270,739	320,512
2037	apc	384,052	245,440	181,366	229,545	273,229	329,258
	att	388,984	249,702	185,023	233,659	277,670	334,050
	fix	392,986	253,265	188,135	237,106	281,345	337,963
2038	apc	410,387	250,212	182,427	234,561	283,335	346,654
	att	416,032	255,016	186,509	239,196	288,374	352,123
	fix	420,334	258,801	189,784	242,854	292,297	356,319
2039	apc	438,510	254,761	183,384	239,588	293,767	364,943
	att	444,920	260,136	187,906	244,772	299,442	371,138
	fix	449,520	264,136	191,335	248,635	303,610	375,616
2040	apc	468,488	259,097	184,249	244,640	304,545	384,166
	att	475,714	265,069	189,225	250,399	310,896	391,134
	fix	480,609	269,278	192,800	254,462	315,306	395,891
Sum 2018	apc	294,490	294,461	293,834	294,394	294,424	294,456
	att	294,488	294,460	293,833	294,393	294,424	294,454
	fix	294,471	294,443	293,816	294,376	294,406	294,438
2019	apc	306,420	306,168	303,620	305,719	305,995	306,273
	att	306,413	306,162	303,613	305,713	305,988	306,267
	fix	306,344	306,092	303,543	305,643	305,918	306,197
2020	apc	319,701	318,822	313,177	317,603	318,549	319,518
	att	319,681	318,801	313,155	317,581	318,527	319,496
	fix	319,524	318,645	313,000	317,425	318,371	319,340
2021	apc	334,407	332,278	322,484	329,860	331,969	334,161
	att	334,357	332,228	322,434	329,809	331,919	334,110
	fix	334,085	331,956	322,163	329,537	331,646	333,838
2022	apc	350,639	346,414	331,544	342,384	346,170	350,159
	att	350,542	346,317	331,449	342,288	346,074	350,062
	fix	350,128	345,903	331,038	341,874	345,659	349,647
2023	apc	368,477	361,089	340,337	355,097	361,090	367,488
	att	368,313	360,925	340,176	354,935	360,926	367,324
	fix	367,732	360,346	339,603	354,356	360,346	366,742
2024	apc	388,003	376,163	348,839	367,944	376,687	386,147
	att	387,747	375,908	348,589	367,690	376,431	385,891
	fix	386,981	375,143	347,835	366,926	375,665	385,122
2025	apc	409,315	391,516	357,040	380,892	392,945	406,161
	att	408,942	391,145	356,679	380,523	392,573	405,786
	fix	407,970	390,176	355,727	379,558	391,602	404,812
2026	apc	432,508	407,028	364,926	393,914	409,849	427,552
	att	431,990	406,514	364,426	393,403	409,334	427,032
	fix	430,799	405,328	363,266	392,222	408,145	425,837
2027	apc	457,688	422,596	372,488	406,992	427,393	450,358
	att	456,996	421,911	371,823	406,310	426,705	449,663
	fix	455,575	420,496	370,445	404,904	425,288	448,234
2028	apc	484,965	438,125	379,720	420,110	445,578	474,613
	att	484,071	437,239	378,865	419,230	444,687	473,713
	fix	482,413	435,590	377,263	417,591	443,033	472,044
2029	apc	514,446	453,527	386,616	433,253	464,393	500,351
	att	513,321	452,413	385,542	432,145	463,272	499,216
	fix	511,421	450,524	383,716	430,270	461,376	497,303
2030	apc	546,252	468,733	393,180	446,415	483,847	527,614
	att	544,869	467,361	391,862	445,053	482,466	526,216

```
        fix          542,726    465,235    389,813    442,942    480,330    524,059
2031 apc             580,480    483,663    399,401    459,575    503,923    556,423
        att          578,814    482,009    397,815    457,933    502,257    554,737
        fix          576,435    479,648    395,548    455,590    499,884    552,339
2032 apc             617,256    498,271    405,291    472,735    524,636    586,822
        att          615,287    496,314    403,418    470,793    522,662    584,829
        fix          612,682    493,724    400,940    468,226    520,059    582,199
2033 apc             656,682    512,501    410,848    485,884    545,974    618,836
        att          654,398    510,221    408,669    483,620    543,676    616,520
        fix          651,577    507,414    405,993    480,839    540,853    613,670
2034 apc             698,891    526,331    416,091    499,032    567,960    652,518
        att          696,290    523,718    413,596    496,437    565,327    649,872
        fix          693,266    520,702    410,732    493,452    562,294    646,817
2035 apc             744,007    539,732    421,033    512,186    590,606    687,908
        att          741,090    536,778    418,213    509,250    587,632    684,937
        fix          737,876    533,562    415,168    506,068    584,398    681,686
2036 apc             792,174    552,708    425,704    525,370    613,948    725,077
        att          788,953    549,409    422,553    522,090    610,634    721,788
        fix          785,560    545,998    419,334    518,716    607,203    718,352
2037 apc             843,509    565,243    430,113    538,590    637,998    764,063
        att          840,002    561,599    426,629    534,965    634,346    760,471
        fix          836,438    557,995    423,239    531,399    630,723    756,860
2038 apc             898,159    577,359    434,298    551,883    662,803    804,941
        att          894,389    573,367    430,477    547,909    658,819    801,070
        fix          890,664    569,572    426,919    544,156    655,007    797,291
2039 apc             956,238    589,059    438,280    565,265    688,391    847,761
        att          952,238    584,722    434,121    560,944    684,083    843,640
        fix          948,364    580,738    430,397    557,004    680,085    839,706
2040 apc           1,017,860    600,361    442,083    578,761    714,793    892,576
        att        1,013,668    595,680    437,583    574,094    710,175    888,240
        fix        1,009,659    591,510    433,698    569,970    705,997    884,166
```

```
> fCtable( addmargins( apply( dmp[,paste(2020+c(-2:0,1:4*5)),,,"att"],
+                             2:4, sum ), 2 ), row.vars=c(2,1), w=9 )
```

```
         imod       apc       att       fix       p20       p40       p60
sex t
M   2018        163,046   163,031   162,695   162,996   163,014   163,031
    2019        169,921   169,787   168,426   169,557   169,713   169,871
    2020        177,504   177,038   174,029   176,421   176,956   177,504
    2025        227,155   217,909   199,718   212,735   219,519   226,953
    2030        299,745   260,187   220,633   249,815   270,791   295,261
    2035        400,956   298,297   236,477   286,589   330,343   384,353
    2040        537,954   330,611   248,358   323,695   399,279   497,106
W   2018        131,442   131,429   131,138   131,397   131,410   131,423
    2019        136,492   136,375   135,187   136,156   136,275   136,396
    2020        142,177   141,763   139,126   141,160   141,571   141,992
    2025        181,787   173,236   156,961   167,788   173,054   178,833
    2030        245,124   207,174   171,229   195,238   211,675   230,955
    2035        340,134   238,481   181,736   222,661   257,289   300,584
    2040        475,714   265,069   189,225   250,399   310,896   391,134
Sum 2018        294,489   294,460   293,833   294,393   294,424   294,455
    2019        306,414   306,162   303,613   305,713   305,989   306,267
    2020        319,680   318,801   313,156   317,581   318,527   319,496
    2025        408,942   391,145   356,679   380,523   392,573   405,786
    2030        544,869   467,362   391,862   445,053   482,466   526,217
    2035        741,090   536,778   418,213   509,250   587,633   684,936
    2040      1,013,668   595,680   437,582   574,094   710,175   888,240
```

For the sake of simplicity we expand the `dmp` to cover the years where we have *observed* number of prevalent cases.

```
> dmp <- dmp[,c(rep(1,22),1:23),,,]
> dmp[,1:22,,,] <- NA
> dimnames( dmp )[[2]] <- paste(1996:2040)
> str( dmp )
 num [1:100, 1:45, 1:2, 1:6, 1:3] NA NA NA NA NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:100] "0" "1" "2" "3" ...
  ..$ t   : chr [1:45] "1996" "1997" "1998" "1999" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
> pr.obs <- aperm(pr.obs[,,,"DM"],c(2,3,1))
> str( pr.obs )
 'table' num [1:100, 1:22, 1:2] 0 3 10 9 9 15 26 32 38 34 ...
 - attr(*, "dimnames")=List of 3
  ..$ A  : chr [1:100] "0" "1" "2" "3" ...
  ..$ P  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ sex: chr [1:2] "M" "W"
> for( i in 1:dim(dmp)[4] )
+ for( j in 1:dim(dmp)[5] )
+ dmp[,1:22,,i,j] <- pr.obs
```

Now `dmp` has the number of diabetes patients in Denmark by age (0–99), calendar time (1996-01-01 – 2040-01-01) and sex. Additionally classified by prediction scenario (6×3), but for dates 1996-01-01 through 2017-01-01 entries are identical across these:

```
> fCtable( addmargins( apply( dmp[,,,,"att"], 2:4, sum ), 2 ),
+          row.vars=c(2,1), w=9 )
          imod       apc       att       fix       p20       p40       p60
sex t
M   1996         42,594    42,594    42,594    42,594    42,594    42,594
    1997         46,465    46,465    46,465    46,465    46,465    46,465
    1998         49,908    49,908    49,908    49,908    49,908    49,908
    1999         53,931    53,931    53,931    53,931    53,931    53,931
    2000         57,780    57,780    57,780    57,780    57,780    57,780
    2001         61,514    61,514    61,514    61,514    61,514    61,514
    2002         65,350    65,350    65,350    65,350    65,350    65,350
    2003         70,167    70,167    70,167    70,167    70,167    70,167
    2004         75,851    75,851    75,851    75,851    75,851    75,851
    2005         81,646    81,646    81,646    81,646    81,646    81,646
    2006         86,227    86,227    86,227    86,227    86,227    86,227
    2007         90,759    90,759    90,759    90,759    90,759    90,759
    2008         95,723    95,723    95,723    95,723    95,723    95,723
    2009        101,854   101,854   101,854   101,854   101,854   101,854
    2010        108,420   108,420   108,420   108,420   108,420   108,420
    2011        115,960   115,960   115,960   115,960   115,960   115,960
    2012        126,936   126,936   126,936   126,936   126,936   126,936
    2013        134,908   134,908   134,908   134,908   134,908   134,908
    2014        140,184   140,184   140,184   140,184   140,184   140,184
    2015        144,880   144,880   144,880   144,880   144,880   144,880
    2016        149,669   149,669   149,669   149,669   149,669   149,669
    2017        154,876   154,876   154,876   154,876   154,876   154,876
    2018        163,046   163,031   162,695   162,996   163,014   163,031
```

	2019	169,921	169,787	168,426	169,557	169,713	169,871
	2020	177,504	177,038	174,029	176,421	176,956	177,504
	2021	185,823	184,699	179,491	183,488	184,680	185,918
	2022	194,917	192,696	184,806	190,694	192,831	195,082
	2023	204,815	200,947	189,956	197,989	201,368	204,975
	2024	215,546	209,372	194,928	205,341	210,266	215,594
	2025	227,155	217,909	199,718	212,735	219,519	226,953
	2026	239,674	226,487	204,310	220,148	229,110	239,056
	2027	253,145	235,050	208,700	227,570	239,033	251,921
	2028	267,616	243,549	212,885	234,993	249,290	265,565
	2029	283,134	251,942	216,863	242,410	259,876	280,006
	2030	299,745	260,187	220,633	249,815	270,791	295,261
	2031	317,494	268,251	224,193	257,201	282,031	311,344
	2032	336,436	276,112	227,552	264,570	293,602	328,276
	2033	356,618	283,743	230,709	271,917	305,502	346,069
	2034	378,106	291,142	233,681	279,255	317,746	364,754
	2035	400,956	298,297	236,477	286,589	330,343	384,353
	2036	425,242	305,217	239,116	293,938	343,318	404,903
	2037	451,018	311,897	241,606	301,306	356,676	426,421
	2038	478,357	318,351	243,968	308,713	370,445	448,947
	2039	507,318	324,586	246,215	316,172	384,641	472,502
	2040	537,954	330,611	248,358	323,695	399,279	497,106
W	1996	40,852	40,852	40,852	40,852	40,852	40,852
	1997	43,876	43,876	43,876	43,876	43,876	43,876
	1998	46,409	46,409	46,409	46,409	46,409	46,409
	1999	49,261	49,261	49,261	49,261	49,261	49,261
	2000	52,353	52,353	52,353	52,353	52,353	52,353
	2001	55,185	55,185	55,185	55,185	55,185	55,185
	2002	57,832	57,832	57,832	57,832	57,832	57,832
	2003	62,252	62,252	62,252	62,252	62,252	62,252
	2004	66,868	66,868	66,868	66,868	66,868	66,868
	2005	71,661	71,661	71,661	71,661	71,661	71,661
	2006	74,960	74,960	74,960	74,960	74,960	74,960
	2007	77,639	77,639	77,639	77,639	77,639	77,639
	2008	81,088	81,088	81,088	81,088	81,088	81,088
	2009	85,459	85,459	85,459	85,459	85,459	85,459
	2010	89,736	89,736	89,736	89,736	89,736	89,736
	2011	94,804	94,804	94,804	94,804	94,804	94,804
	2012	104,267	104,267	104,267	104,267	104,267	104,267
	2013	110,608	110,608	110,608	110,608	110,608	110,608
	2014	114,737	114,737	114,737	114,737	114,737	114,737
	2015	118,133	118,133	118,133	118,133	118,133	118,133
	2016	121,595	121,595	121,595	121,595	121,595	121,595
	2017	125,162	125,162	125,162	125,162	125,162	125,162
	2018	131,442	131,429	131,138	131,397	131,410	131,423
	2019	136,492	136,375	135,187	136,156	136,275	136,396
	2020	142,177	141,763	139,126	141,160	141,571	141,992
	2021	148,534	147,529	142,943	146,321	147,239	148,192
	2022	155,625	153,621	146,643	151,594	153,243	154,980
	2023	163,498	159,978	150,220	156,946	159,558	162,349
	2024	172,201	166,536	153,661	162,349	166,165	170,297
	2025	181,787	173,236	156,961	167,788	173,054	178,833
	2026	192,316	180,027	160,116	173,255	180,224	187,976
	2027	203,851	186,861	163,123	178,740	187,672	197,742
	2028	216,455	193,690	165,980	184,237	195,397	208,148
	2029	230,187	200,471	168,679	189,735	203,396	219,210
	2030	245,124	207,174	171,229	195,238	211,675	230,955

	2031	261,320	213,758	173,622	200,732	220,226	243,393
	2032	278,851	220,202	175,866	206,223	229,060	256,553
	2033	297,780	226,478	177,960	211,703	238,174	270,451
	2034	318,184	232,576	179,915	217,182	247,581	285,118
	2035	340,134	238,481	181,736	222,661	257,289	300,584
	2036	363,711	244,192	183,437	228,152	267,316	316,885
	2037	388,984	249,702	185,023	233,659	277,670	334,050
	2038	416,032	255,016	186,509	239,196	288,374	352,123
	2039	444,920	260,136	187,906	244,772	299,442	371,138
	2040	475,714	265,069	189,225	250,399	310,896	391,134
Sum	1996	83,446	83,446	83,446	83,446	83,446	83,446
	1997	90,341	90,341	90,341	90,341	90,341	90,341
	1998	96,317	96,317	96,317	96,317	96,317	96,317
	1999	103,192	103,192	103,192	103,192	103,192	103,192
	2000	110,133	110,133	110,133	110,133	110,133	110,133
	2001	116,699	116,699	116,699	116,699	116,699	116,699
	2002	123,182	123,182	123,182	123,182	123,182	123,182
	2003	132,419	132,419	132,419	132,419	132,419	132,419
	2004	142,719	142,719	142,719	142,719	142,719	142,719
	2005	153,307	153,307	153,307	153,307	153,307	153,307
	2006	161,187	161,187	161,187	161,187	161,187	161,187
	2007	168,398	168,398	168,398	168,398	168,398	168,398
	2008	176,811	176,811	176,811	176,811	176,811	176,811
	2009	187,313	187,313	187,313	187,313	187,313	187,313
	2010	198,156	198,156	198,156	198,156	198,156	198,156
	2011	210,764	210,764	210,764	210,764	210,764	210,764
	2012	231,203	231,203	231,203	231,203	231,203	231,203
	2013	245,516	245,516	245,516	245,516	245,516	245,516
	2014	254,921	254,921	254,921	254,921	254,921	254,921
	2015	263,013	263,013	263,013	263,013	263,013	263,013
	2016	271,264	271,264	271,264	271,264	271,264	271,264
	2017	280,038	280,038	280,038	280,038	280,038	280,038
	2018	294,489	294,460	293,833	294,393	294,424	294,455
	2019	306,414	306,162	303,613	305,713	305,989	306,267
	2020	319,680	318,801	313,156	317,581	318,527	319,496
	2021	334,357	332,227	322,434	329,809	331,918	334,110
	2022	350,542	346,317	331,449	342,288	346,074	350,062
	2023	368,313	360,925	340,176	354,935	360,926	367,324
	2024	387,748	375,908	348,589	367,690	376,432	385,891
	2025	408,942	391,145	356,679	380,523	392,573	405,786
	2026	431,990	406,514	364,426	393,403	409,333	427,032
	2027	456,997	421,910	371,823	406,310	426,705	449,662
	2028	484,071	437,239	378,864	419,230	444,687	473,713
	2029	513,322	452,413	385,542	432,146	463,272	499,216
	2030	544,869	467,362	391,862	445,053	482,466	526,217
	2031	578,814	482,009	397,815	457,933	502,258	554,737
	2032	615,287	496,314	403,417	470,792	522,662	584,828
	2033	654,398	510,222	408,669	483,620	543,676	616,519
	2034	696,289	523,717	413,596	496,437	565,327	649,872
	2035	741,090	536,778	418,213	509,250	587,633	684,936
	2036	788,953	549,409	422,553	522,090	610,633	721,787
	2037	840,002	561,599	426,629	534,964	634,346	760,472
	2038	894,389	573,367	430,478	547,909	658,819	801,070
	2039	952,238	584,723	434,120	560,944	684,084	843,639
	2040	1,013,668	595,680	437,582	574,094	710,175	888,240

We would like to see the overall change in the number of diabetes patients, as recorded in the structure `dmp`, both for the total population but also separatelu for men and women.

```
> DMall <- dmp[,,"M",,] + dmp[,,"W",,]
> str( DMall )
 num [1:100, 1:45, 1:6, 1:3] 0 6 13 17 18 30 44 56 57 68 ...
 - attr(*, "dimnames")=List of 4
  ..$ a   : chr [1:100] "0" "1" "2" "3" ...
  ..$ t   : chr [1:45] "1996" "1997" "1998" "1999" ...
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
> DMcum <- apply( DMall, 2:4, cumsum )
> DMcum <- DMcum[c(1,1:100),,,]
> DMcum[1,,,] <- 0
> DMcum <- DMcum
> str( DMcum )
 num [1:101, 1:45, 1:6, 1:3] 0 0 6 19 36 54 84 128 184 241 ...
 - attr(*, "dimnames")=List of 4
  ..$ a   : chr [1:101] "0" "0" "1" "2" ...
  ..$ t   : chr [1:45] "1996" "1997" "1998" "1999" ...
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
```

Finally we can plot the predicted numbers from the different scenarios (stacking across age-classes):

```
> ryr <- c(1996:2040,2040:1996)
> leg <- c("DM Incidence fixed at 2016 level",
+          "Linear projection of DM inc. from 2016",
+          "Attenuated linear projection of DM inc.",
+          "DM inc. increasing 2.0%/y",
+          "DM inc. increasing 4.0%/y",
+          "DM inc. increasing 6.0%/y")
> names( leg ) <- c("fix","apc","att","p20","p40","p60")
> cbind( leg )
    leg
fix "DM Incidence fixed at 2016 level"
apc "Linear projection of DM inc. from 2016"
att "Attenuated linear projection of DM inc."
p20 "DM inc. increasing 2.0%/y"
p40 "DM inc. increasing 4.0%/y"
p60 "DM inc. increasing 6.0%/y"
> pl.num <-
+ function( wh.m, mtxt, xl=c(1996,2030), ymax=800 )
+ {
+ par( mfrow=c(3,2), mar=c(2,1,0,3), oma=c(0,0,2,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ ltno <- 0
+ for( wh.i in c("apc","fix","att","p20","p40","p60") ) # c(2,3,1,4:6) )
+    {
+ ltno <- ltno+1
+ plot( NA,
+       xlim=xl, xlab="", xaxt="n", xaxs="i",
+       ylim=c(0,ymax), xaxs="i", yaxs="i", yaxt="n", ylab="" )
+ for( i in 1:10 ) polygon( ryr, c( DMcum[1+(i-1)*10,,wh.i,wh.m],
+                                   rev( DMcum[1+ i   *10,,wh.i,wh.m] ) )/1000,
+                           col=gray( (14-i)/15 ), border=gray(0.8) ) #"transparent" )
```

```
+ abline( h=seq(50,400,50), v=seq(1990,2040,5), col=gray(1), lty="14", lend="butt" )
+ abline( v=2017, lty=3 )
+ for( i in seq(55,85,10) ) text( xl[2]-0.2, DMcum[paste(i),paste(xl[2]-1),wh.i,wh.m]/1000,
+                                 paste( i-5,"-",i+4,sep=""), adj=c(1,1),
+                                 col=gray(0.95), font=2 )
+ axis( side=4, at=0:9*100 )
+ axis( side=4, at=seq(0,9,1/4)*100, labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1990,2040,10) )
+ axis( side=1, at=seq(1995,2040,5), labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1995,2040,1), labels=NA, tcl=-0.2 )
+ text( xl[1]+2, ymax*0.95, letters[ltno], font=2, adj=c(1,0), cex=1.4 )
+ text( xl[1]+3, ymax*0.95, paste( #wh.i, ": ",
+                                  leg[wh.i], sep="" ), adj=c(0,0) )
+    }
+ mtext( mtxt, side=3, line=1, outer=TRUE, cex=0.66, adj=0 )
+ mtext("No.(1000s)", side=3, line=0.5, outer=TRUE, cex=0.66, adj=1, at=0.49 )
+ mtext("No.(1000s)", side=3, line=0.5, outer=TRUE, cex=0.66, adj=1, at=0.99 )
+ }
> pl.num( "apc", "Mortality constantly decreasing", ymax=600, xl=c(1996,2030) )


> pl.num( "apc", "Mortality constantly decreasing", ymax=600, xl=c(1996,2040) )


> pl.num( "att", "Mortality decrease attenuated", ymax=600, xl=c(1996,2030) )


> pl.num( "att", "Mortality decrease attenuated", ymax=600, xl=c(1996,2040) )


> pl.num( "fix", "Mortality fixed at 2016", ymax=600, xl=c(1996,2030) )


> pl.num( "fix", "Mortality fixed at 2016", ymax=600, xl=c(1996,2040) )
```

From figure 10.3 it appears that it is the prediction of incidence rates of diabetes that carries the major differences of more than 100,000 patients in 2040. The decrease in the number of incident cases during the period 2012–2014and the pick-up during 2015, make any type of prediction pretty hap-hazard.

Thus the prediction of the number of future patients is crucially dependent on the tiny amount of information available about future diabetes incidence rates in the rather odd behaviour of the rates in the years 2011 through 2016.

10.5.1 Analyses by sex

Her we replicate (some of) the plots above, separately for each sex

```
> DMm <- dmp[,,"M",,]
> DMf <- dmp[,,"W",,]
> DMcm <- apply( DMm, 2:4, cumsum )
> DMcf <- apply( DMf, 2:4, cumsum )
> DMcm <- DMcm[c(1,1:100),,,]
> DMcf <- DMcf[c(1,1:100),,,]
> DMcm[1,,,] <- DMcf[1,,,] <- 0
> str( DMcm )
```

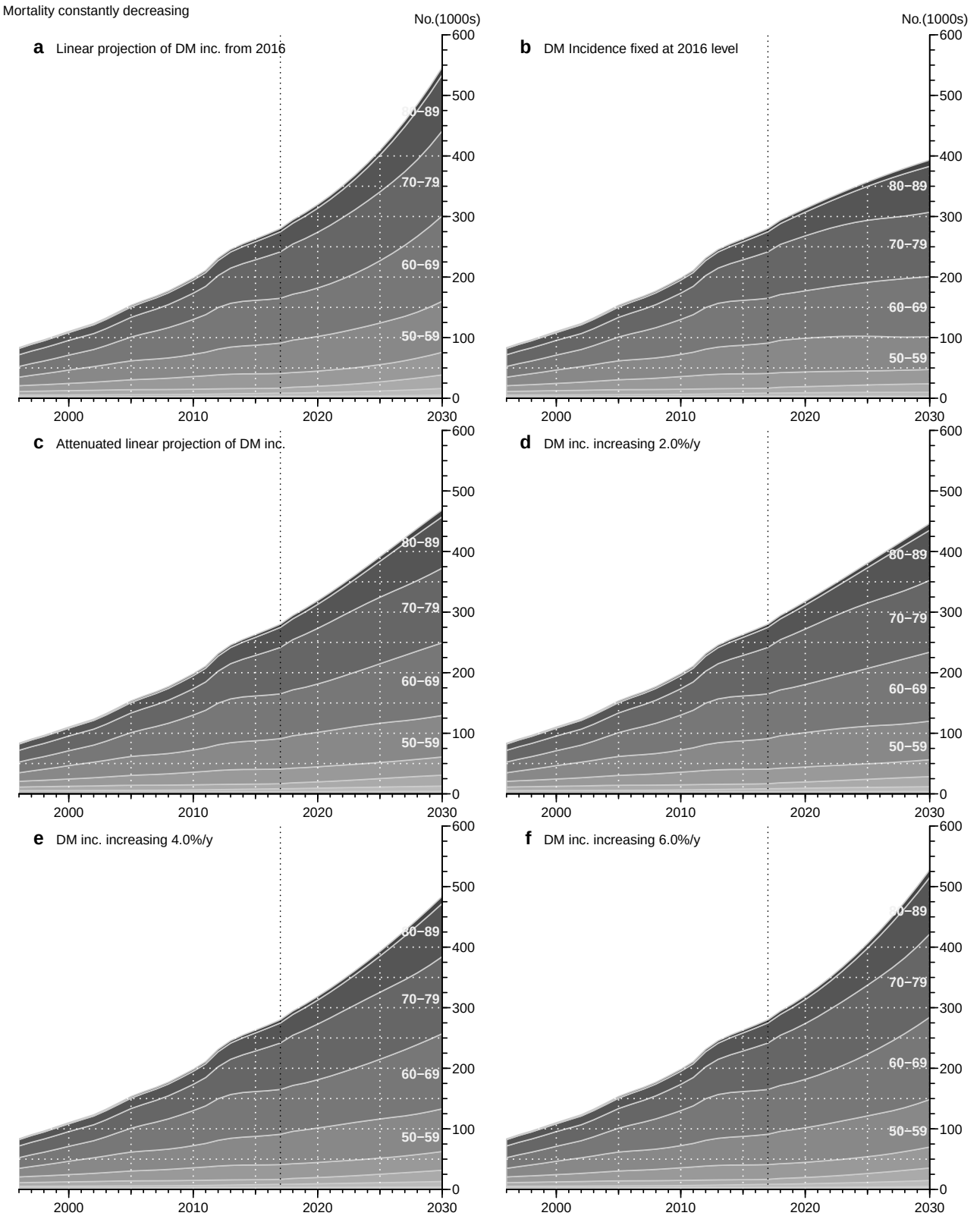


Figure 10.3: *Predicted number of diabetes patients in Denmark under different incidence rate scenarios, using a continuing decrease in mortality (both for non-DM and DM persons).* `./graph/pred-prnum-apc`

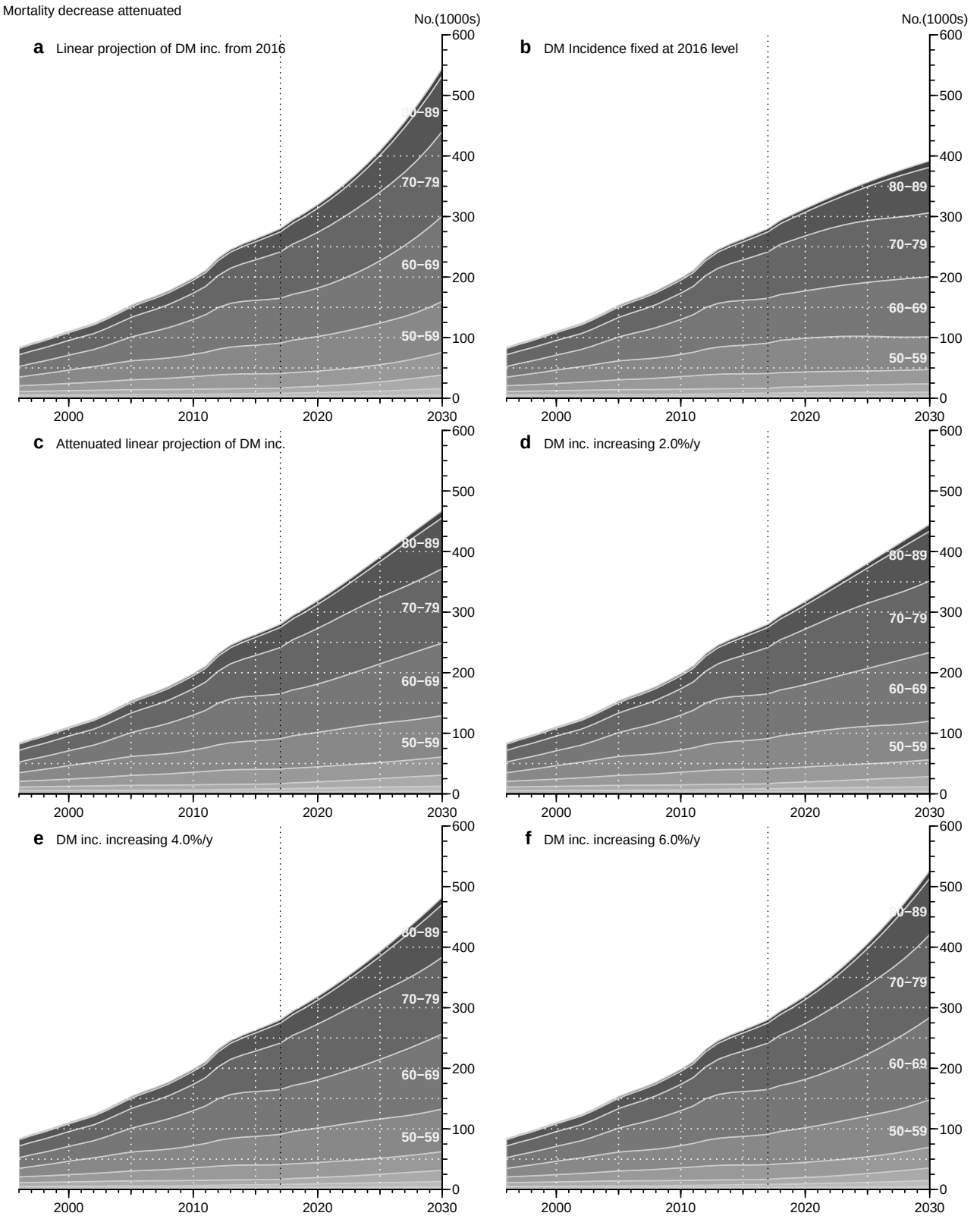


Figure 10.4: *Predicted number of diabetes patients in Denmark under different incidence rate scenarios, using an attenuated decrease in mortality (both for non-DM and DM persons).* `./graph/pred-prnum-att`

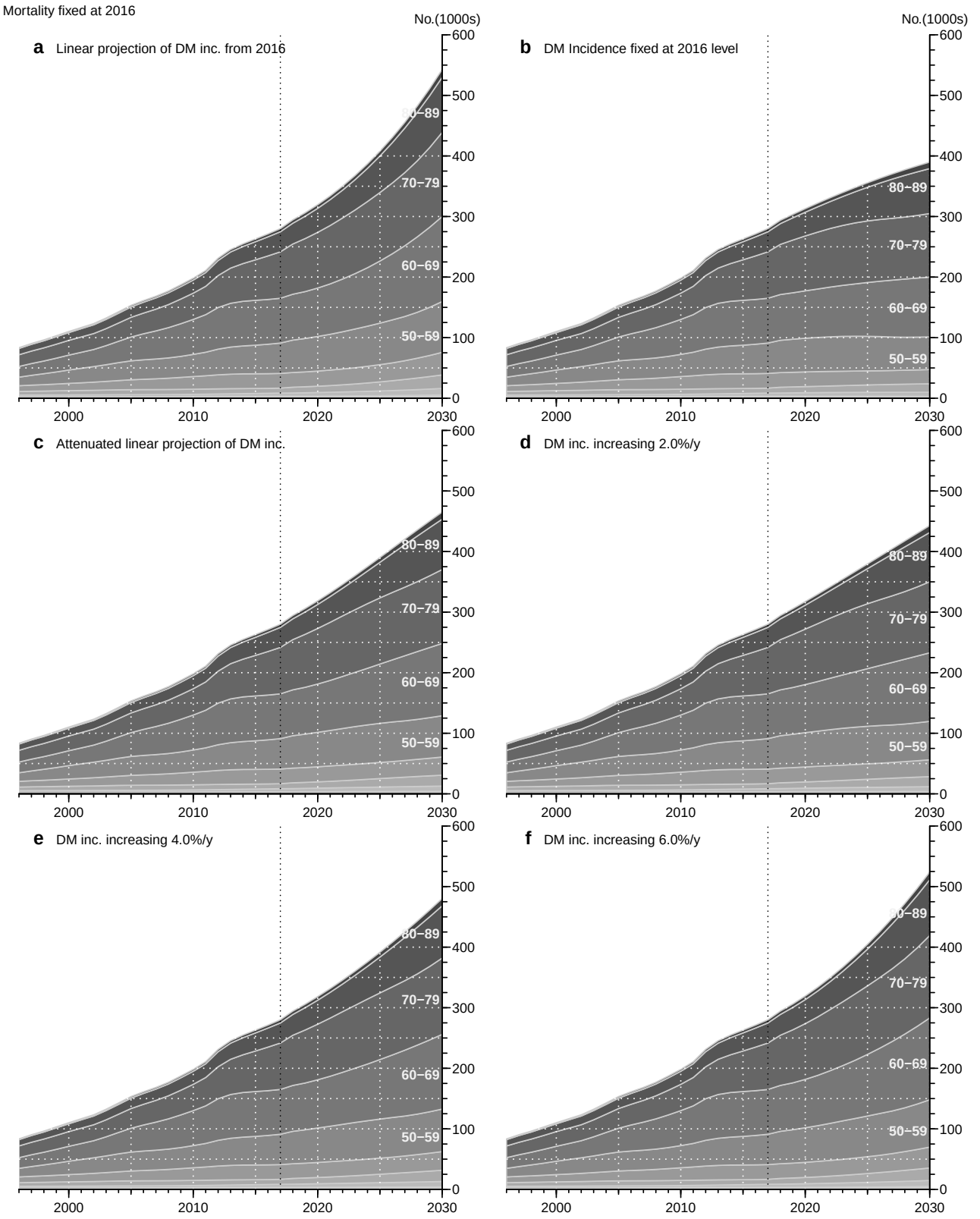


Figure 10.5: *Predicted number of diabetes patients in Denmark under different incidence rate scenarios, using mortality (both for non-DM and DM persons) fixed at the 2017 level.* `./graph/pred-prnum-fix`

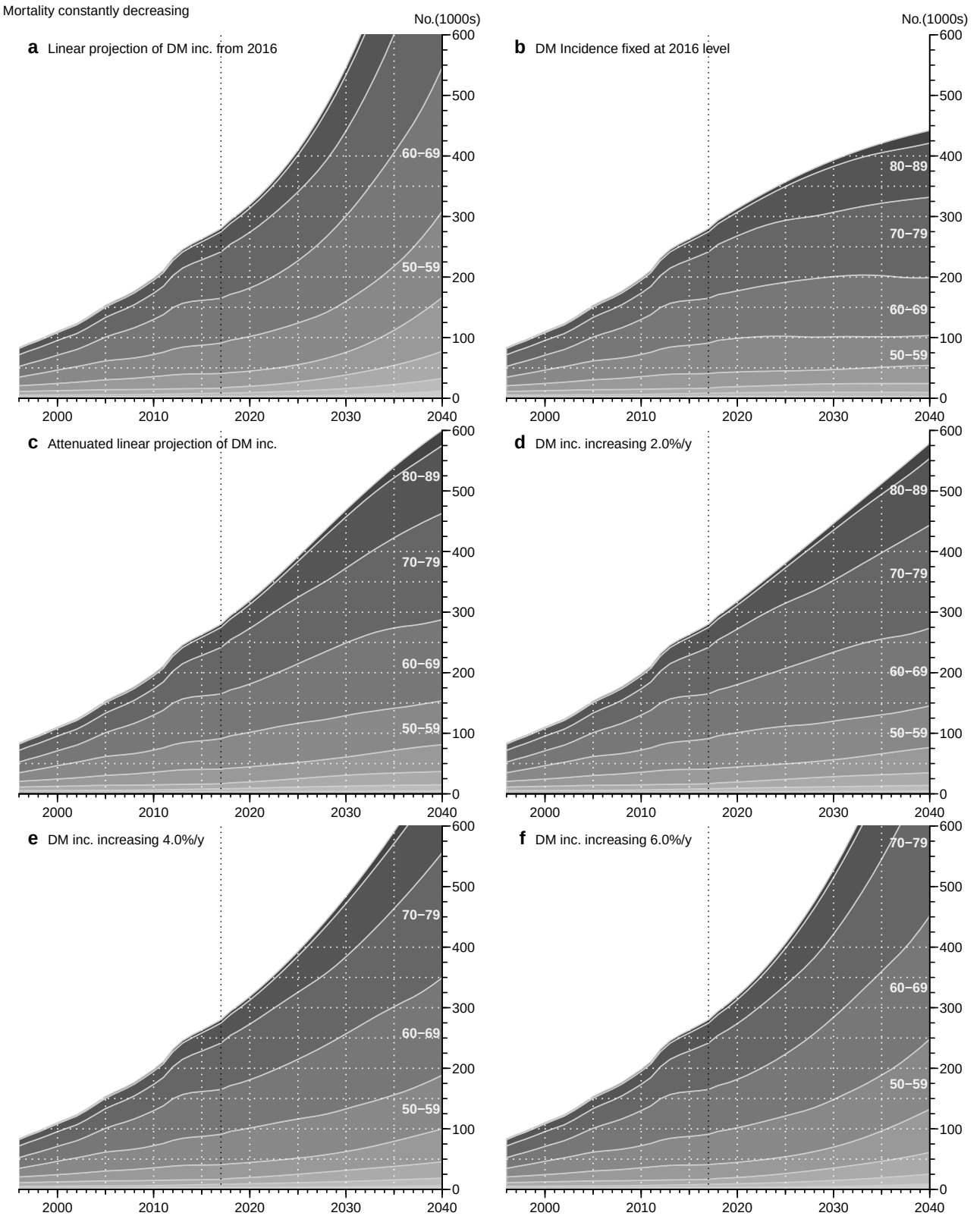


Figure 10.6: *Predicted number of diabetes patients in Denmark under different incidence rate scenarios, using a continuing decrease in mortality (both for non-DM and DM persons).* `./graph/pred-prnum-apc2040`

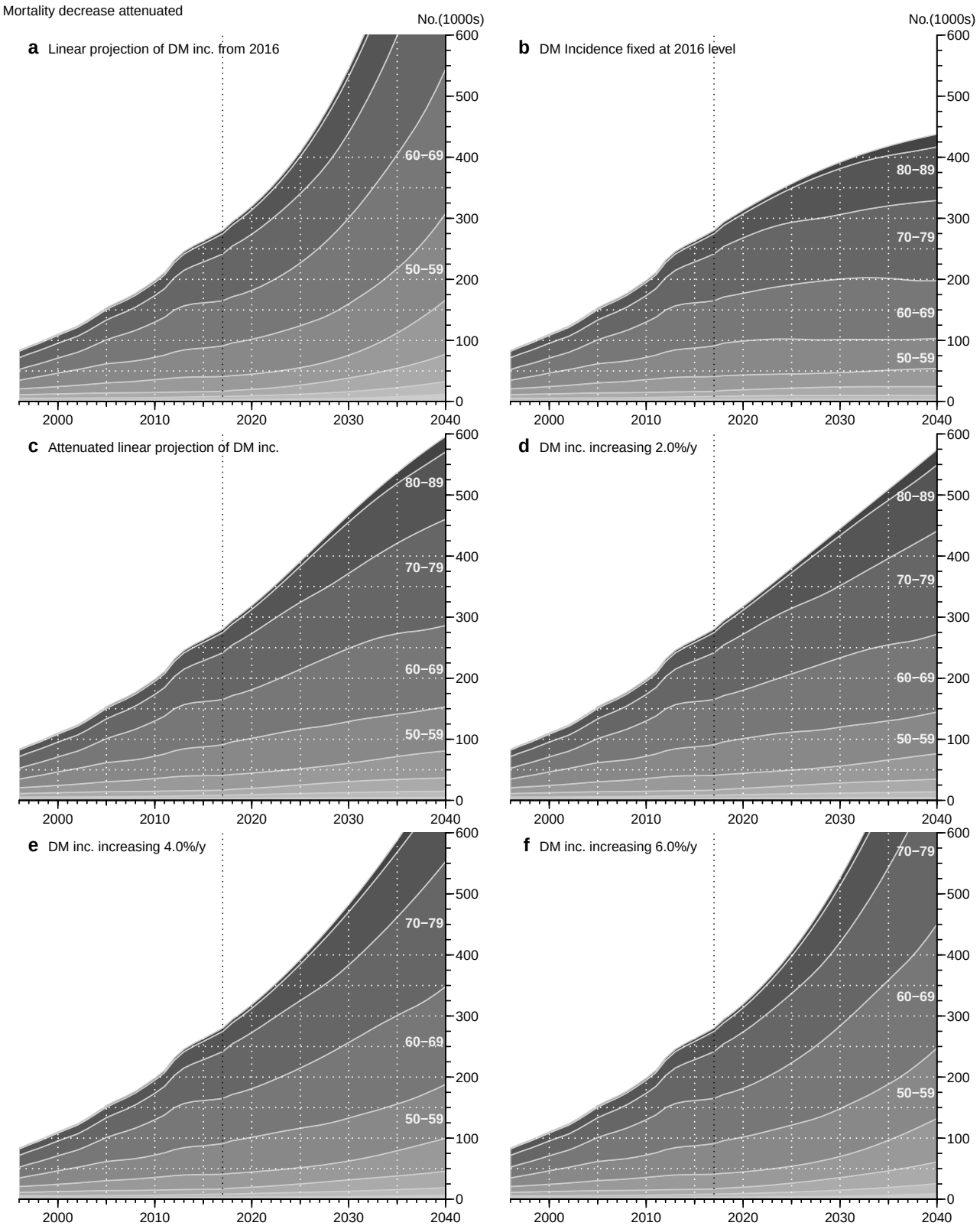


Figure 10.7: *Predicted number of diabetes patients in Denmark under different incidence rate scenarios, using an attenuated decrease in mortality (both for non-DM and DM persons).* `./graph/pred-prnum-att2040`

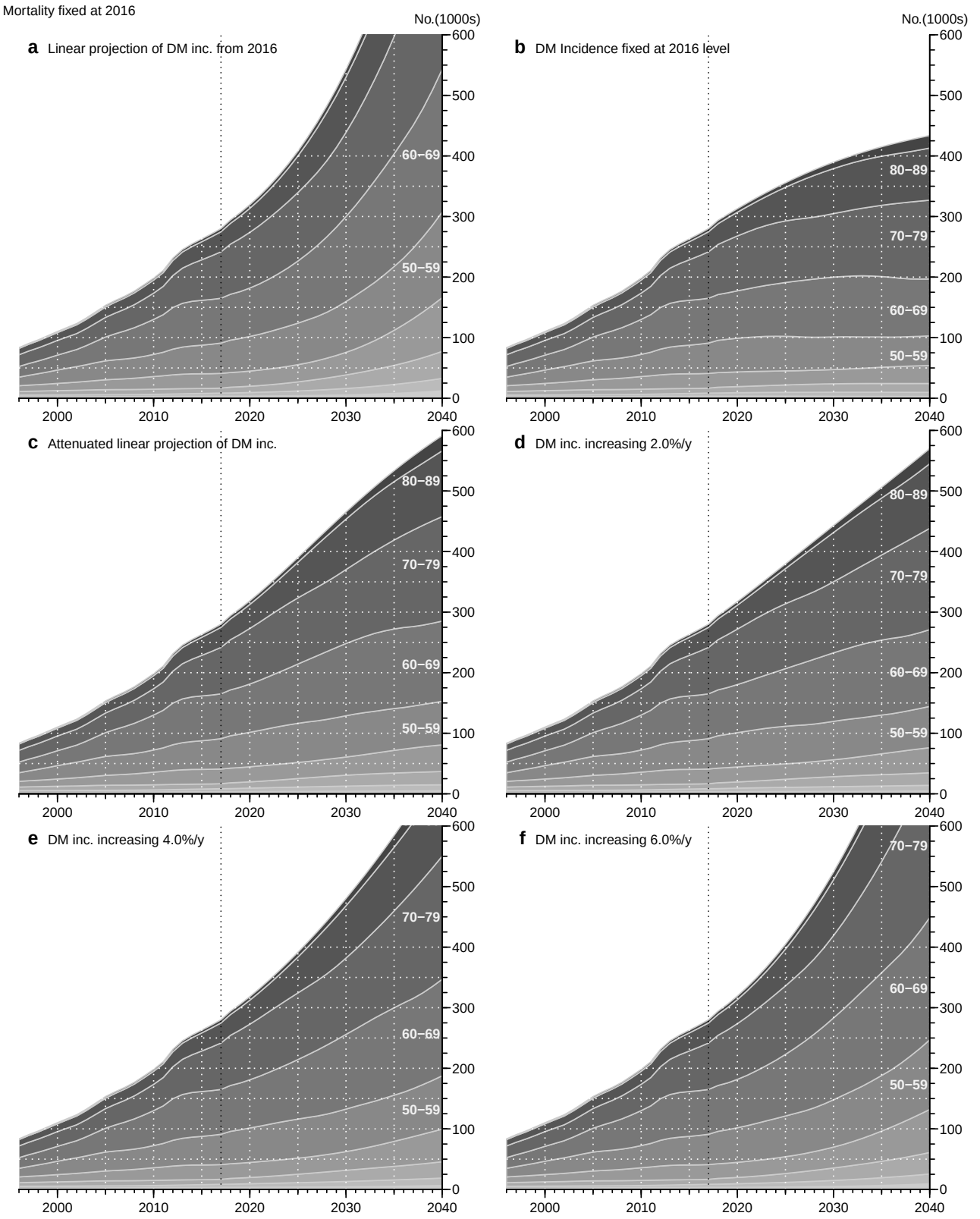


Figure 10.8: *Predicted number of diabetes patients in Denmark under different incidence rate scenarios, using mortality (both for non-DM and DM persons) fixed at the 2017 level.* `./graph/pred-prnum-fix2040`

```
 num [1:101, 1:45, 1:6, 1:3] 0 0 3 13 22 31 46 72 104 142 ...
 - attr(*, "dimnames")=List of 4
  ..$ a   : chr [1:101] "0" "0" "1" "2" ...
  ..$ t   : chr [1:45] "1996" "1997" "1998" "1999" ...
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
> str( DMcf )
 num [1:101, 1:45, 1:6, 1:3] 0 0 3 6 14 23 38 56 80 99 ...
 - attr(*, "dimnames")=List of 4
  ..$ a   : chr [1:101] "0" "0" "1" "2" ...
  ..$ t   : chr [1:45] "1996" "1997" "1998" "1999" ...
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"
```

Besides plotting the cumulative nombers by age, we also want to plot the *relative distribution* over age. This is basically the same as dividing by the cumulative numbers at 100 years by a sweep:

```
> DMpm <- sweep( DMcm, 2:4, DMcm[101,,,], "/" )
> DMpf <- sweep( DMcf, 2:4, DMcf[101,,,], "/" )
> DMpa <- sweep( DMcum, 2:4, DMcum[101,,,], "/" )
```

The plots we will show are therefore just the same as aove but with slightly modified code; whe need a bit different colouring and axes:

```
> library(RColorBrewer)
> clb <- brewer.pal(9,"Blues")[c(1,1:9)]
> clr <- brewer.pal(9,"Reds" )[c(1,1:9)]
> clg <- brewer.pal(9,"Greys")[c(1,1:9)]
> xl <- c(1996,2030)
> plrel <- function( both=TRUE )
+    {
+ par( mfcol=c(if(both) 3 else 2,2), mar=c(2,1,0,3), oma=c(0,0,2,0),
+      mgp=c(3,1,0)/1.6, las=1, bty="n" )
+
+ ymax <- 275
+
+ plot( NA,
+       xlim=xl, xlab="", xaxt="n", xaxs="i",
+       ylim=c(0,ymax), xaxs="i", yaxs="i", yaxt="n", ylab="" )
+ for( i in 1:10 ) polygon( ryr, c( DMcm[1+(i-1)*10,,"att","att"],
+                                   rev( DMcm[1+ i   *10,,"att","att"] ) )/1000,
+                           col=clb[i], border="transparent" )
+ abline( h=seq(50,400,50), v=seq(1990,2040,5), col=gray(1), lty="14", lend="butt" )
+ abline( v=2017, lty=1 )
+ for( i in seq(55,85,10) ) text( xl[2]-0.2, DMcm[paste(i),paste(xl[2]-1),"att","att"]/1000,
+                                 paste( i-5,"-",i+4,sep=""), adj=c(1,1),
+                                 col=gray(0.95), font=2 )
+ axis( side=4, at=0:9*50 )
+ axis( side=4, at=seq(0,9,1/4)*100, labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1990,2040,10) )
+ axis( side=1, at=seq(1995,2040,5), labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1995,2040,1), labels=NA, tcl=-0.2 )
+ text( xl[1]+2, ymax*0.98, "a", font=2, adj=c(1,1), cex=1.4 )
+
+ plot( NA,
```

```
+       xlim=xl, xlab="", xaxt="n", xaxs="i",
+       ylim=c(0,ymax), xaxs="i", yaxs="i", yaxt="n", ylab="" )
+ for( i in 1:10 ) polygon( ryr, c( DMcf[1+(i-1)*10,,"att","att"],
+                              rev( DMcf[1+ i   *10,,"att","att"] ) )/1000,
+                            col=clr[i], border="transparent" )
+ abline( h=seq(50,400,50), v=seq(1990,2040,5), col=gray(1), lty="14", lend="butt" )
+ abline( v=2017, lty=1 )
+ for( i in seq(55,85,10) ) text( xl[2]-0.2, DMcf[paste(i),paste(xl[2]-1),"att","att"]/1000,
+                                 paste( i-5,"-",i+4,sep=""), adj=c(1,1),
+                                 col=gray(0.95), font=2 )
+ axis( side=4, at=0:9*50 )
+ axis( side=4, at=seq(0,9,1/4)*100, labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1990,2040,10) )
+ axis( side=1, at=seq(1995,2040,5), labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1995,2040,1), labels=NA, tcl=-0.2 )
+ text( xl[1]+2, ymax*0.98, "c", font=2, adj=c(1,1), cex=1.4 )
+
+ if( both ){
+ ymax <- 500
+ plot( NA,
+       xlim=xl, xlab="", xaxt="n", xaxs="i",
+       ylim=c(0,ymax), xaxs="i", yaxs="i", yaxt="n", ylab="" )
+ for( i in 1:10 ) polygon( ryr, c( DMcum[1+(i-1)*10,,"att","att"],
+                              rev( DMcum[1+ i   *10,,"att","att"] ) )/1000,
+                            col=clg[i], border="transparent" )
+ abline( h=seq(50,400,50), v=seq(1990,2040,5), col=gray(1), lty="14", lend="butt" )
+ abline( v=2017, lty=1 )
+ for( i in seq(55,85,10) ) text( xl[2]-0.2, DMcum[paste(i),paste(xl[2]-1),"att","att"]/1000,
+                                 paste( i-5,"-",i+4,sep=""), adj=c(1,1),
+                                 col=gray(0.95), font=2 )
+ axis( side=4, at=0:9*100 )
+ axis( side=4, at=seq(0,9,1/4)*100, labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1990,2040,10) )
+ axis( side=1, at=seq(1995,2040,5), labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1995,2040,1), labels=NA, tcl=-0.2 )
+ text( xl[1]+2, ymax*0.98, "e", font=2, adj=c(1,1), cex=1.4 )
+         }
+ # Relative distribution
+
+ ymax <- 100
+ plot( NA,
+       xlim=xl, xlab="", xaxt="n", xaxs="i",
+       ylim=c(0,ymax), xaxs="i", yaxs="i", yaxt="n", ylab="" )
+ for( i in 1:10 ) polygon( ryr, c( DMpm[1+(i-1)*10,,"att","att"],
+                              rev( DMpm[1+ i   *10,,"att","att"] ) )*100,
+                            col=clb[i], border="transparent" )
+ for( i in seq(55,85,10) ) text( xl[2]-0.2, DMpm[paste(i),paste(xl[2]-1),"att","att"]*100,
+                                 paste( i-5,"-",i+4,sep=""), adj=c(1,1),
+                                 col=gray(0.95), font=2 )
+ abline( h=1:9*10, v=seq(1990,2040,5), col=gray(1), lty="14", lend="butt" )
+ abline( v=2017, lty=1 )
+ axis( side=4, at=0:5*20 )
+ axis( side=4, at=seq(0,9,1/10)*100, labels=NA, tcl=-0.5 )
+ axis( side=4, at=seq(0,100,5), labels=NA, tcl=-0.4 )
+ axis( side=4, at=seq(0,100,1), labels=NA, tcl=-0.2 )
+ axis( side=1, at=seq(1990,2040,10) )
+ axis( side=1, at=seq(1995,2040,5), labels=NA, tcl=-0.3 )
```

```
+ axis( side=1, at=seq(1995,2040,1), labels=NA, tcl=-0.2 )
+ text( xl[1]+2, ymax*0.98, "b", font=2, adj=c(1,1), cex=1.4, col="white" )
+
+ plot( NA,
+       xlim=xl, xlab="", xaxt="n", xaxs="i",
+       ylim=c(0,ymax), xaxs="i", yaxs="i", yaxt="n", ylab="" )
+ for( i in 1:10 ) polygon( ryr, c( DMpf[1+(i-1)*10,,"att","att"],
+                              rev( DMpf[1+ i   *10,,"att","att"] ) )*100,
+                           col=clr[i], border="transparent" )
+ for( i in seq(55,85,10) ) text( xl[2]-0.2, DMpf[paste(i),paste(xl[2]-1),"att","att"]*100,
+                                 paste( i-5,"-",i+4,sep=""), adj=c(1,1),
+                                 col=gray(0.95), font=2 )
+ abline( h=1:9*10, v=seq(1990,2040,5), col=gray(1), lty="14", lend="butt" )
+ abline( v=2017, lty=1 )
+ axis( side=4, at=0:5*20 )
+ axis( side=4, at=seq(0,9,1/10)*100, labels=NA, tcl=-0.5 )
+ axis( side=4, at=seq(0,100,5), labels=NA, tcl=-0.4 )
+ axis( side=4, at=seq(0,100,1), labels=NA, tcl=-0.2 )
+ axis( side=1, at=seq(1990,2040,10) )
+ axis( side=1, at=seq(1995,2040,5), labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1995,2040,1), labels=NA, tcl=-0.2 )
+ text( xl[1]+2, ymax*0.98, "d", font=2, adj=c(1,1), cex=1.4, col="white" )
+
+ if( both ){
+ plot( NA,
+       xlim=xl, xlab="", xaxt="n", xaxs="i",
+       ylim=c(0,ymax), xaxs="i", yaxs="i", yaxt="n", ylab="" )
+ for( i in 1:10 ) polygon( ryr, c( DMpa[1+(i-1)*10,,"att","att"],
+                              rev( DMpa[1+ i   *10,,"att","att"] ) )*100,
+                           col=clg[i], border="transparent" )
+ for( i in seq(55,85,10) ) text( xl[2]-0.2, DMpa[paste(i),paste(xl[2]-1),"att","att"]*100,
+                                 paste( i-5,"-",i+4,sep=""), adj=c(1,1),
+                                 col=gray(0.95), font=2 )
+ abline( h=1:9*10, v=seq(1990,2040,5), col=gray(1), lty="14", lend="butt" )
+ abline( v=2017, lty=1 )
+ axis( side=4, at=0:5*20 )
+ axis( side=4, at=seq(0,9,1/10)*100, labels=NA, tcl=-0.5 )
+ axis( side=4, at=seq(0,100,5), labels=NA, tcl=-0.4 )
+ axis( side=4, at=seq(0,100,1), labels=NA, tcl=-0.2 )
+ axis( side=1, at=seq(1990,2040,10) )
+ axis( side=1, at=seq(1995,2040,5), labels=NA, tcl=-0.3 )
+ axis( side=1, at=seq(1995,2040,1), labels=NA, tcl=-0.2 )
+ text( xl[1]+2, ymax*0.98, "f", font=2, adj=c(1,1), cex=1.4, col="white" )
+       }
+ mtext("No.(1000s)", side=3, line=0.5, outer=TRUE, adj=1, at=0.49 )
+ mtext("%", side=3, line=0.5, outer=TRUE, adj=1, at=0.97 )
+     }
> plrel()

> plrel(FALSE)
```

10.6 Time trends in old age patients

Here we compute the fraction of patienst that are 70, resp, 80 or older:

```
> str( dmp )
 num [1:100, 1:45, 1:2, 1:6, 1:3] 0 3 10 9 9 15 26 32 38 34 ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:100] "0" "1" "2" "3" ...
  ..$ t   : chr [1:45] "1996" "1997" "1998" "1999" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ imod: chr [1:6] "apc" "att" "fix" "p20" ...
  ..$ mmod: chr [1:3] "apc" "att" "fix"

> t70 <- addmargins( apply( dmp[71:100,,,"att","att"], 2:3, sum ), 2 )
> t80 <- addmargins( apply( dmp[81:100,,,"att","att"], 2:3, sum ), 2 )
>  t0 <- addmargins( apply( dmp[      ,,,"att","att"], 2:3, sum ), 2 )
> ( wh.yr <- paste(2017+c(0:2,3+0:4*5)) )

[1] "2017" "2018" "2019" "2020" "2025" "2030" "2035" "2040"

> fCtable( cbind( t70, t80, t0 )[wh.yr,] )

              M          W        Sum          M          W        Sum          M          W        Su

2017     60,008     54,975    114,983     16,974     21,347     38,321    154,876    125,162    280,03
2018     64,501     58,240    122,741     17,832     22,012     39,844    163,031    131,429    294,46
2019     68,999     61,214    130,213     19,445     23,208     42,653    169,787    136,375    306,16
2020     73,281     64,129    137,411     21,147     24,416     45,563    177,038    141,763    318,80
2025     96,214     80,298    176,512     33,687     33,156     66,843    217,909    173,236    391,14
2030    120,210     98,240    218,450     50,417     45,178     95,595    260,187    207,174    467,36
2035    145,557    118,132    263,690     61,914     53,818    115,732    298,297    238,481    536,77
2040    171,033    138,380    309,412     73,135     62,184    135,319    330,611    265,069    595,68

> round( cbind( t70/t0, t80/t0 )[wh.yr,] * 100, 1 )

        M    W  Sum    M    W  Sum
2017 38.7 43.9 41.1 11.0 17.1 13.7
2018 39.6 44.3 41.7 10.9 16.7 13.5
2019 40.6 44.9 42.5 11.5 17.0 13.9
2020 41.4 45.2 43.1 11.9 17.2 14.3
2025 44.2 46.4 45.1 15.5 19.1 17.1
2030 46.2 47.4 46.7 19.4 21.8 20.5
2035 48.8 49.5 49.1 20.8 22.6 21.6
2040 51.7 52.2 51.9 22.1 23.5 22.7

> rbind( cbind( round(t70), round(t70/t0*100, 1) )[wh.yr,],
+        cbind( round(t80), round(t80/t0*100, 1) )[wh.yr,] )

          M      W    Sum    M    W  Sum
2017  60008  54975 114983 38.7 43.9 41.1
2018  64501  58240 122741 39.6 44.3 41.7
2019  68999  61214 130213 40.6 44.9 42.5
2020  73281  64129 137411 41.4 45.2 43.1
2025  96214  80298 176512 44.2 46.4 45.1
2030 120210  98240 218450 46.2 47.4 46.7
2035 145557 118132 263690 48.8 49.5 49.1
2040 171033 138380 309412 51.7 52.2 51.9
2017  16974  21347  38321 11.0 17.1 13.7
2018  17832  22012  39844 10.9 16.7 13.5
2019  19445  23208  42653 11.5 17.0 13.9
2020  21147  24416  45563 11.9 17.2 14.3
2025  33687  33156  66843 15.5 19.1 17.1
2030  50417  45178  95595 19.4 21.8 20.5
2035  61914  53818 115732 20.8 22.6 21.6
2040  73135  62184 135319 22.1 23.5 22.7
```

We would also like to have the relative fraction of women:

```
> round( sweep( t0[,1:2], 1, t0[,"Sum"], "/" )*100, 1 )[wh.yr,]
      sex
t          M    W
  2017 55.3 44.7
  2018 55.4 44.6
  2019 55.5 44.5
  2020 55.5 44.5
  2025 55.7 44.3
  2030 55.7 44.3
  2035 55.6 44.4
  2040 55.5 44.5


> elapsed()
----------------------
2019-01-11 at 18:11:14
Time elapsed: 00:01:17
----------------------
```

. . . now input from `comp.tex`

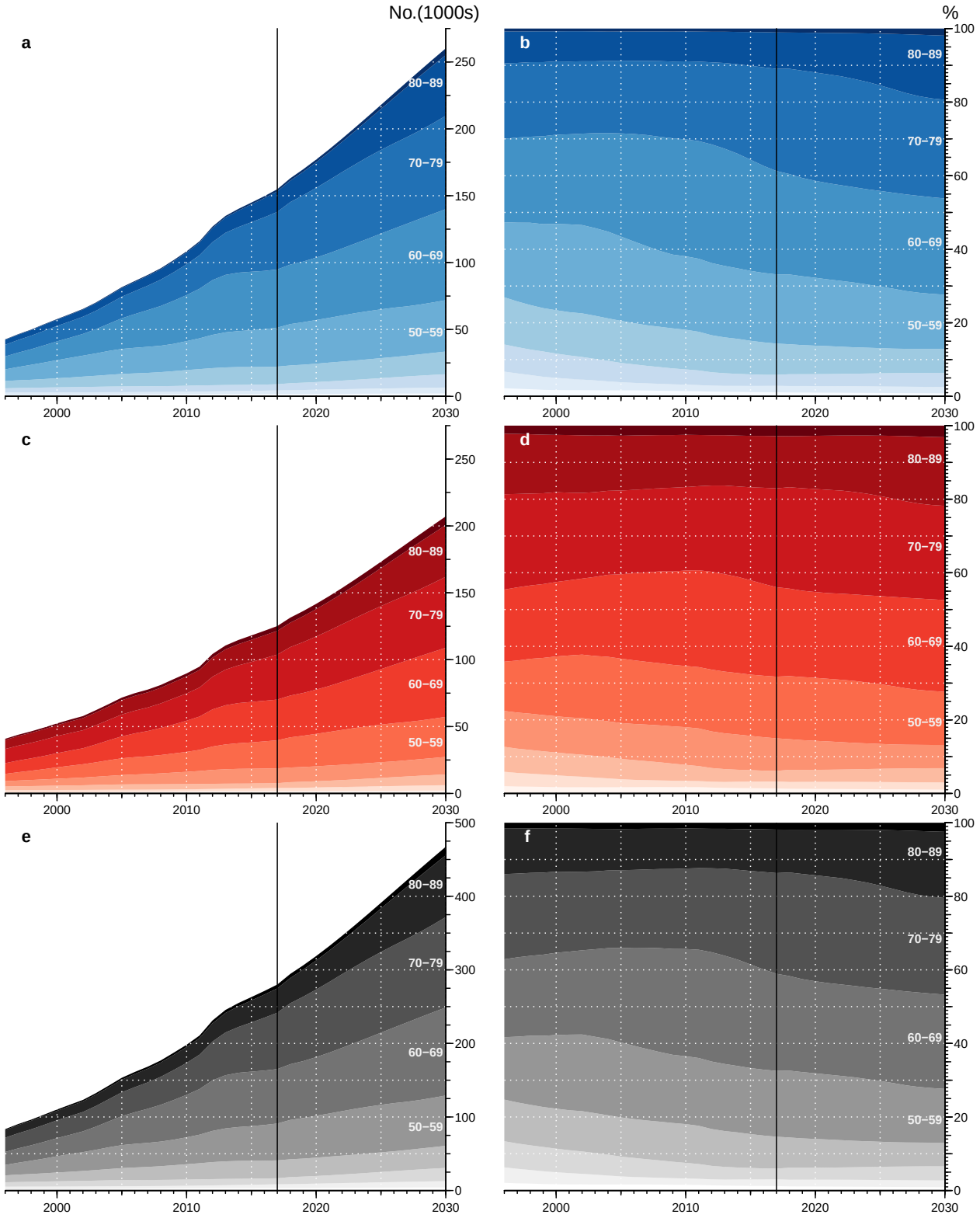


Figure 10.9: *Observed and predicted number of diabetes patients in Denmark 1996–2030. Left panels are absolute numbers, right panels show the relative distribution in 10-year age-classes. Blue is men, red women and gray both sexes together. The vertical black line indicates the end of data and start of projection.* `./graph/pred-cumulate-sex-a`

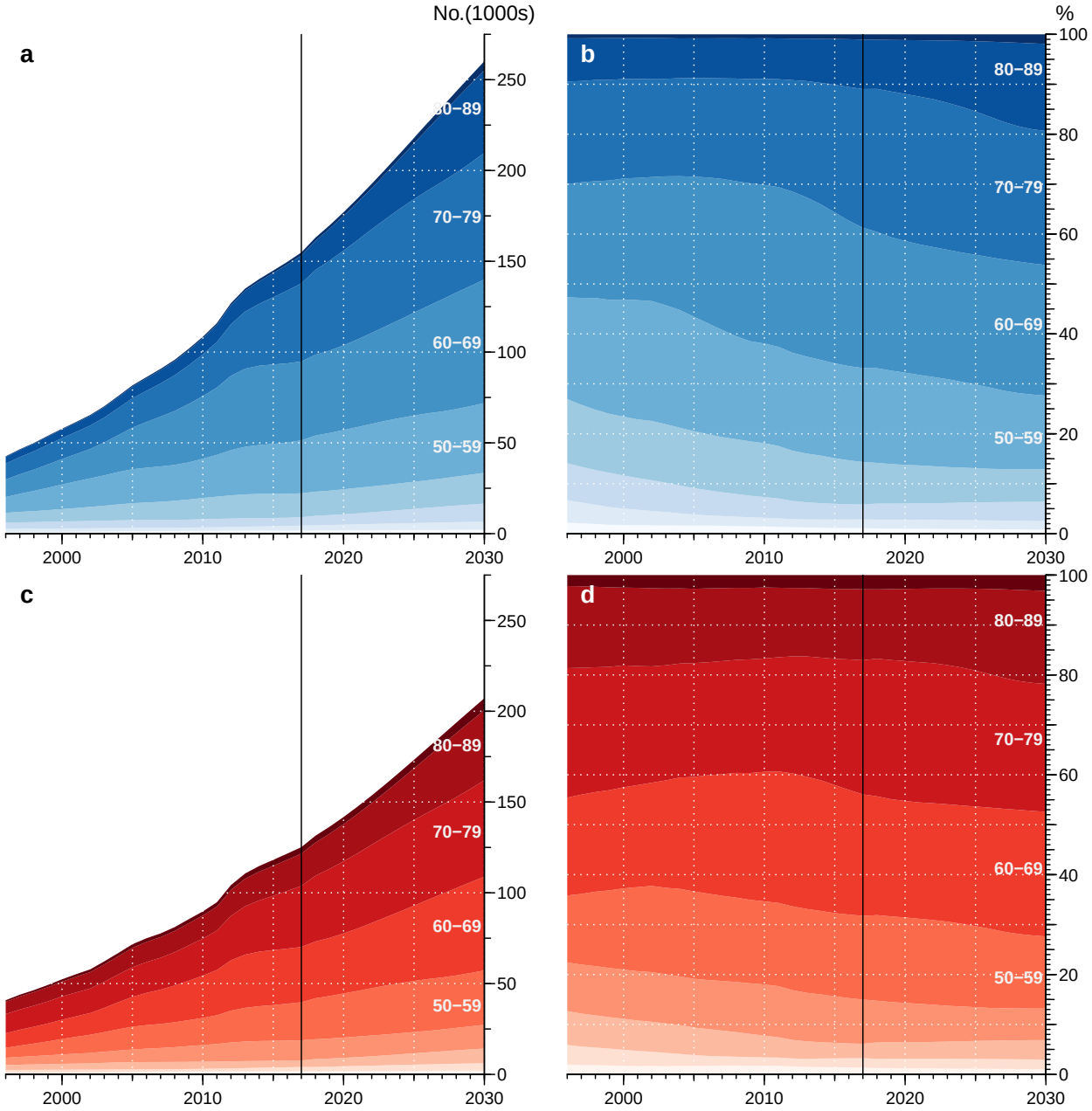


Figure 10.10: *Observed and predicted number of diabetes patients in Denmark 1996–2030. Left panels are absolute numbers, right panels show the relative distribution in 10-year age-classes. Blue is men, red women. The vertical black line indicates the end of data and start of projection.*
`./graph/pred-cumulate-sex`

Chapter 11

Components of prevalence

The purpose of this chapter is to use the estimated transition rates to predict the prevalences during follow-up based on prevalences at 1996 and incidence and mortality rates in the period 1996–2016. This is not *per se* an interesting endeavour, because we have the prevalence data available, but it will serve as an illustration that the rates are adequately modelled and that the degree of approximation is adequate when using an interval length as chosen.

Moreover the methodology used for the forecasting of rates and prevalences in previous chapter where we used different scenarios can also be applied here, now instead using different so-called counterfactual scenarios.

So first we load the data:

```
> library( Epi )
> start()
---------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-01-06 12:36:31
---------------------------------------
> load( file="../nydata/inits.Rda")
> load( file="../nydata/allrates.Rda")
> load( file="../nydata/prevalences.Rda")
> lls()
   name     mode     class          dim                  size(Kb)
1  a.pt     numeric  numeric        1200                      9.4
2  akn      list     matrix         3 3                       8.1
3  cstr     function function       1                        12.2
4  d.pt     numeric  numeric        240                       1.9
5  elapsed  function function       1                        11.5
6  fC       function function       1                         2.7
7  fCp      function function       1                         2.8
8  fCtable  function function       1                         2.8
9  ini.time numeric  POSIXct POSIXt 1                         0.3
10 int      numeric  numeric        1                         0.1
11 Lambda   numeric  array          1200 528 2 9         89,238.5
12 mod      list     matrix         3 3                   6,155.2
13 Mu.DM    numeric  array          1200 528 2 9         89,238.5
14 Mu.nD    numeric  array          1200 528 2 9         89,238.5
15 nk.a     numeric  numeric        1                         0.1
16 nk.c     numeric  numeric        1                         0.1
17 nk.d     numeric  numeric        1                         0.1
18 nk.p     numeric  numeric        1                         0.1
19 p.pt     numeric  numeric        528                       4.2
```

```
20 parr     numeric  array          2 3 3 201 22 3       1,882.0
21 pr.fin   numeric  array          3 3 1200 3             348.9
22 pr.ini   numeric  array          3 3 1200 3             348.9
23 pr.obs   numeric  xtabs table    2 100 22 3             114.1
24 qn       function function       1                       16.8
25 start    function function       1                        6.9
26 t.pt     numeric  numeric        529                      4.2
```

We shall use the simulation scheme to predict the course of DM prevalence development in the population under various scenarios of mortality and incidence development. So we set up structures to hold results and clarify calculations:

`pr.fit` — array of empirical age-specific prevalences at 1996-01-01–1.1.2017-01-01, smoothed by natural splines.

`TR` — array of transition probabilities between states Well and DM and Death. Transition probabilities are computed under the 4 different scenarios combining mortality and incidence rates either as they actually developed 1996–2016 or assuming they were constant at the 1996 level. These refer to intervals of length `int` and are therefore labeled on the period dimension by the midpoint of these, a total of 21/`int`.

`prprv` — array of `pr`edicted `pr`evalences based on the initial prevalences at 1996-01-01 and the transition probabilities as put in `TR`. The scenario dimension refers to the 4 scenarios: "`obs`", "`m-fix`", "`i-fix`" and "`all-f`", but this dimension in the array is expanded by 3 extra levels "`mort`", "`inc`" and "`const`" that are to be filled with the part of the prevalences that are attributable to decrease in mortality, increase in incidence and the disequilibrium between rates and prevalence in 1996. Likewise, the period dimension is expanded by one level relative to that in `TR` (replacing `p.pt` (period points) with `t.pt` (time points), since this dimansion refer to points in time and not time intervals (periods).

`prn` — array of predicted *number* of DM patients in one-year age classes at the 1 January each year. So the same structure as `prprv`, but with substantially fewer entries.

11.1 Transition probabilities

In order to get the predicted *number* of persons by age, period and prediction type, we need the (1-step) transition matrices at all combinations of age (a) and date (p), this is put in array.

Note that the structures `Lambda`, `Mu.nD` and `Mu.DM`, contain predicted rates up to 2040 under different scenarios, that we do not need:

```
> str( Lambda )
 num [1:1200, 1:528, 1:2, 1:9] 6.14e-06 6.18e-06 6.22e-06 6.26e-06 6.30e-06 ...
 - attr(*, "dimnames")=List of 4
  ..$ a  : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" ...
  ..$ p  : chr [1:528] "1996.04166666667" "1996.125" "1996.20833333333" "1996.29166666667" ...
  ..$ sex: chr [1:2] "M" "W"
  ..$ mod: chr [1:9] "ap" "apc" "gam" "LCa" ...
```

. . . we only need the dates till 2017-01-01, and the models `ap` and `apc`:

```
> # 2nd dimension of rates is all the way to 2040, only need to the end
> # of 2016
> dimnames(Lambda)[[2]][252+0:2]
[1] "2016.95833333333" "2017.04166666667" "2017.125"
> Lambda <- Lambda[,1:252,,1:2]
> Mu.nD  <- Mu.nD [,1:252,,1:2]
> Mu.DM  <- Mu.DM [,1:252,,1:2]
> states <- c("nD","DM")
> TR <- NArray( c( dimnames(Lambda),
+                  list( from = states,
+                          to = states,
+                       scene = c("obs","m-fix","i-fix","all-f" ) ) ) )
> str( TR )
 logi [1:1200, 1:252, 1:2, 1:2, 1:2, 1:2, 1:4] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 7
  ..$ a    : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ p    : chr [1:252] "1996.04166666667" "1996.125" "1996.20833333333" "1996.29166666667" .
  ..$ sex  : chr [1:2] "M" "W"
  ..$ mod  : chr [1:2] "ap" "apc"
  ..$ from : chr [1:2] "nD" "DM"
  ..$ to   : chr [1:2] "nD" "DM"
  ..$ scene: chr [1:4] "obs" "m-fix" "i-fix" "all-f"
```

The situation where both the mortality rates and incidence rates are fixed at the 1996 level is trivial, because transition probabilities in that case only depend on age and not on period.

In order to fill `TR`, we need the cumulative incidences over intervals of length `int`. But these were exactly the ones we predicted in the previous sections by setting the person years equal to `int` in the data frame supplied to the `newdata` argument.

So we can now compute the one-`int`-step transition matrices for every combination of `a.pt` and `p.pt`, both in steps of `int` (in this case 0.083 year):

```
> int
[1] 0.08333333
> TR[,,,,"nD","nD","obs"] <- 1 - Lambda - Mu.nD
> TR[,,,,"nD","DM","obs"] <-     Lambda
> TR[,,,,"DM","nD","obs"] <- 0
> TR[,,,,"DM","DM","obs"] <- 1 - Mu.DM
```

Note that we have not included the "Dead" state in the calculations, because we only bother about the *fraction* of diabetes patients in each age class at each time-point. So the probabilities we compute do not sum to 1 within the "from" states; we only compute the fraction of the persons alive that end up being alive at the next time point.

When we fix the mortality or incidence at the 1996 level we just replace the expressions above with expressions where we replace the date dimension by `rep(1,np)`, (where `np` is the number of periods) for either incidence, mortality or both:

```
> ( np <- dim(Lambda)[2] )
  p
252
> #
> TR[,,,,"nD","nD","m-fix"] <- 1 - Lambda - Mu.nD[,rep(1,np),,]
> TR[,,,,"nD","DM","m-fix"] <-     Lambda
> TR[,,,,"DM","nD","m-fix"] <- 0
```

```
> TR[,,,,"DM","DM","m-fix"] <- 1 -          Mu.DM[,rep(1,np),,]
> #
> TR[,,,,"nD","nD","i-fix"] <- 1 - Lambda[,rep(1,np),,] - Mu.nD
> TR[,,,,"nD","DM","i-fix"] <-     Lambda[,rep(1,np),,]
> TR[,,,,"DM","nD","i-fix"] <- 0
> TR[,,,,"DM","DM","i-fix"] <- 1 -                         Mu.DM
> #
> TR[,,,,"nD","nD","all-f"] <- 1 - Lambda[,rep(1,np),,] - Mu.nD[,rep(1,np),,]
> TR[,,,,"nD","DM","all-f"] <-     Lambda[,rep(1,np),,]
> TR[,,,,"DM","nD","all-f"] <- 0
> TR[,,,,"DM","DM","all-f"] <- 1 -                         Mu.DM[,rep(1,np),,]
```

We have now collected the transition probabilities between "Well" and "DM" as well as the probabilities of remaining in each of these, all referring to a duration of `int`:

```
> attr( TR, "label" ) <- "Transition probabilities 1996-2016"
> str( TR )
 num [1:1200, 1:252, 1:2, 1:2, 1:2, 1:2, 1:4] 1 1 1 1 1 ...
 - attr(*, "dimnames")=List of 7
  ..$ a    : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ p    : chr [1:252] "1996.04166666667" "1996.125" "1996.20833333333" "1996.29166666667" .
  ..$ sex  : chr [1:2] "M" "W"
  ..$ mod  : chr [1:2] "ap" "apc"
  ..$ from : chr [1:2] "nD" "DM"
  ..$ to   : chr [1:2] "nD" "DM"
  ..$ scene: chr [1:4] "obs" "m-fix" "i-fix" "all-f"
 - attr(*, "label")= chr "Transition probabilities 1996-2016"
> fCp( length( TR ) )
[1] 19,353,600
> save( TR, file="../nydata/TRc.Rda" )
```

11.2 Prediction of the observed prevalences

Note that we do not need to predict the population size; we can get away with only predicting the prevalences as fractions. When we multiply the fraction of persons in states (Well,DM) with the transition matrix, we get fraction of the persons in the previous state that are in states (Well,DM), which does not sum to 1 (because of the ones dying), so we must rescale to prevalence age in each step.

First we set up an array to hold the predicted prevalences under different scenarios. Note that while `Lambda` (and the other rate-arrays) are classified by the *midpoints* of the (a, p) sets of the Lexis diagram (in this case 1 month $\times$ 1 month intervals), we want the prevalences at the *midpoint* of the age-classes, but at the *border* of the periods, thus we need an extra level for the calendar time dimension.

Later we shall also compute the fraction of the prevalences that are attributable to trends in mortality and incidence as well as to the non-stationarity of the rates/prevalences as of 1996, so we put in three extra levels of the last dimension, and one extra levels of the period dimension because we want to predict to the end of the last period too (or, to put it differently, we need an extra first level to hold the starting prevalences as of 1996-01-01).

```
> dpr <- c( dimnames(Lambda)[1:4],
+           list( c(dimnames(TR)[["scene"]],"mort","inc","const","org") ) )
> names( dpr )[c(2,5)] <- c("t","what")
> t.pt[252+0:2]
[1] 2016.917 2017.000 2017.083
> dpr[["t"]] <- t.pt[1:253]
> prprv <- NArray(dpr)
```

In order to update the prevalences at 1996-01-01 to subsequent dates, we must first initialize the prevalences as modeled at 1996-01-01, as well as the prevalences at ages 0 (the new-born — set to 0); note that we exploit the column major storage of arrays to get `prprv` filled with identical values across the two last dimensions (model and scenario).

```
> str( pr.ini )
 num [1:3, 1:3, 1:1200, 1:3] 3.44e-04 5.83e-06 5.37e-04 2.98e-04 4.47e-06 ...
 - attr(*, "dimnames")=List of 4
  ..$ typ: chr [1:3] "T1" "T2" "DM"
  ..$ sex: chr [1:3] "M" "W" "B"
  ..$ A  : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" .
  ..$    : chr [1:3] "Est" "lo" "hi"
> str( prprv )
 logi [1:1200, 1:253, 1:2, 1:2, 1:8] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ t   : chr [1:253] "1996" "1996.08333333333" "1996.16666666667" "1996.25" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ mod : chr [1:2] "ap" "apc"
  ..$ what: chr [1:8] "obs" "m-fix" "i-fix" "all-f" ...
> # Smoothed prevalences at 1996-01-01 - the starting values
> for( sx in c("M","W") ) prprv[,"1996",sx,,] <- pr.ini["DM",sx,,"Est"]
> # Prevalences at age 0
> prprv[1,,"M",,] <- 0
> prprv[1,,"W",,] <- 0
> # check we got values in the right spots
> round( ftable( prprv[1:3,1:3,,,], row.vars=4:1 )*100, 1 )
                                        what obs m-fix i-fix all-f mort inc const org
mod sex t                 a
ap  M   1996              0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                   0.1   0.1   0.1   0.1  0.1 0.1   0.1 0.1
                          0.208333333333333       0.1   0.1   0.1   0.1  0.1 0.1   0.1 0.1
        1996.08333333333  0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                    NA    NA    NA    NA   NA  NA    NA  NA
                          0.208333333333333        NA    NA    NA    NA   NA  NA    NA  NA
        1996.16666666667  0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                    NA    NA    NA    NA   NA  NA    NA  NA
                          0.208333333333333        NA    NA    NA    NA   NA  NA    NA  NA
    W   1996              0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                   0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.208333333333333       0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
        1996.08333333333  0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                    NA    NA    NA    NA   NA  NA    NA  NA
                          0.208333333333333        NA    NA    NA    NA   NA  NA    NA  NA
        1996.16666666667  0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                    NA    NA    NA    NA   NA  NA    NA  NA
                          0.208333333333333        NA    NA    NA    NA   NA  NA    NA  NA
```

```
apc M   1996              0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                   0.1   0.1   0.1   0.1  0.1 0.1   0.1 0.1
                          0.208333333333333       0.1   0.1   0.1   0.1  0.1 0.1   0.1 0.1
        1996.08333333333 0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                    NA    NA    NA    NA   NA  NA    NA  NA
                          0.208333333333333        NA    NA    NA    NA   NA  NA    NA  NA
        1996.16666666667 0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                    NA    NA    NA    NA   NA  NA    NA  NA
                          0.208333333333333        NA    NA    NA    NA   NA  NA    NA  NA
    W   1996              0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                   0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.208333333333333       0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
        1996.08333333333 0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                    NA    NA    NA    NA   NA  NA    NA  NA
                          0.208333333333333        NA    NA    NA    NA   NA  NA    NA  NA
        1996.16666666667 0.0416666666666667      0.0   0.0   0.0   0.0  0.0 0.0   0.0 0.0
                          0.125                    NA    NA    NA    NA   NA  NA    NA  NA
                          0.208333333333333        NA    NA    NA    NA   NA  NA    NA  NA
```

So now we have checked that we have put initial values correctly into `prprv`, basically at the period edge (at 1996) and the nidpoint of the first age category (at 0). Then we can compute the predicted prevalences under the different scenarios. We take the fraction of the population in age class `ia` at time `ip` that end up as diabetes patients at time `ip+1` (and hence in age class `ia+1`), and divide by the fraction of all that remain alive, which is the diabetes patients (survivors and new), *plus* those who survive free of diabetes:

```
> system.time(
+ for( ip in 1:(dim(prprv)[2]-1) )
+ for( ia in 1:(dim(prprv)[1]-1) )
+ prprv[ia+1,ip+1,,,1:4] <-
+   (    prprv[ia,ip,,,1:4]  * TR[ia,ip,,,"DM","DM",]
+    +(1-prprv[ia,ip,,,1:4]) * TR[ia,ip,,,"nD","DM",] ) /
+   (    prprv[ia,ip,,,1:4]  * TR[ia,ip,,,"DM","DM",]
+    +(1-prprv[ia,ip,,,1:4]) * TR[ia,ip,,,"nD","DM",]
+    +(1-prprv[ia,ip,,,1:4]) * TR[ia,ip,,,"nD","nD",] )
+             )
   user  system elapsed
  15.68    0.02   15.68
```

Note that the reason that the last dimension, `scene`, is explicitly mentioned in the array `prprv` is because the length of this dimension in `prprv` is 7, but the corresponding in `TR` only 4 — recall that `prprv` also has thee extra levels to provide for the estimated part of the prevalences attributable to mortality change, incidence changes, and non-equilibrium at 1996.

Moreover, note that the code above is particularly simple because we only need to compute the prevalence at the next date and age. If we had had a more elaborate model with, say complications states, the calculations in the loop would have been a matrix-multiplication updating the state-distribution, but this simplification would have been at the expense of another three loop-levels, namely over the three last dimensions of the `prprv` array — the matrix machinery only operates on 2-dimensional structures, matrices. (Well, maybe some `apply` construction could be set up).

We can then show a few of the predicted prevalences in (%):

```
> round( prprv[900+1:4,1+(0:3)*24,1,1,1,drop=F]*100, 3 )
```

```
, , sex = M, mod = ap, what = obs

                   t
a                    1996  1998  2000  2002
  75.0416666666667 5.729 6.655 7.548 8.567
  75.125           5.740 6.663 7.554 8.573
  75.2083333333333 5.752 6.671 7.561 8.579
  75.2916666666667 5.763 6.679 7.567 8.585


> save( a.pt, prprv, file="../nydata/prprv.Rda" )
> load(               file="../nydata/prprv.Rda" )
```

11.2.1 Checking the prediction

With this initial prediction in place we can now check whether we have made a reasonable approximation to the observed prevalences at 2016-01-01.

In the array `prprv` are all the prevalences as predicted from the prevalence in 1996 using the estimated incidences and mortalities; predicted at intervals of `inc` whereas we have the smoothed empirical prevalences at 1 January 1996,...,2016 in the array `parr`:

```
> str( parr )
 num [1:2, 1:3, 1:3, 1:201, 1:22, 1:3] 3.42e-04 1.26e-04 5.78e-06 6.62e-06 5.35e-04 ...
 - attr(*, "dimnames")=List of 6
  ..$ mod: chr [1:2] "glm" "gam"
  ..$ typ: chr [1:3] "T1" "T2" "DM"
  ..$ sex: chr [1:3] "M" "W" "B"
  ..$ A  : chr [1:201] "0" "0.5" "1" "1.5" ...
  ..$ T  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$    : chr [1:3] "Est" "lo" "hi"
> str( prprv )
 num [1:1200, 1:253, 1:2, 1:2, 1:8] 0 0.000541 0.000545 0.000549 0.000553 ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ t   : chr [1:253] "1996" "1996.08333333333" "1996.16666666667" "1996.25" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ mod : chr [1:2] "ap" "apc"
  ..$ what: chr [1:8] "obs" "m-fix" "i-fix" "all-f" ...
```

Thus we have the predicted age-specific prevalences for men in say 2003 in `prprv[,"2003","M","apc","obs"]`, and the smoothed empirical prevalences (from the prevalence chapter) in `parr["glm","DM","M"■"2003","Est"]`. We now plot these for select years in the same plot framw:

```
> ( wh <- paste(seq(2017,1996,-4)) )
[1] "2017" "2013" "2009" "2005" "2001" "1997"
> pchk <- function( ap=TRUE )
+    {
+ par( mfrow=c(1,2), mar=c(0,0,0,0), oma=c(3,4,1,1), mgp=c(3,1,0)/1.6,
+      las=1, bty="n", lend="butt" )
+ for( sx in c("M","W") )
+    {
+ cl <- ifelse( sx=="M", "blue", "red" )
```

```
+ matplot( as.numeric(dimnames(parr)[[4]]), parr["glm","DM",sx,,wh,"Est"]*100,
+          xlim=c(10,95), ylim=c(0,20), yaxs="i", yaxt="n",
+          xlab="Age", ylab="Prevalence (%)",
+          type="l", col=cl, lty=1, lwd=2 )
+ axis( side=1, at=1:9*10, labels=NA )
+ axis( side=1, at=seq(10,95,5), labels=NA, tcl=-0.3 )
+ if( sx =="M" ){
+ axis( side=2 )
+ axis( side=2, at=1:20*10, labels=NA, tcl=-0.3 ) }
+ matlines( as.numeric(dimnames(prprv)[[1]]), prprv[,wh,sx,"apc","obs"]*100,
+          type="l", col=cl, lty="11", lwd=3, lend="butt" )
+ if( ap )
+ matlines( as.numeric(dimnames(prprv)[[1]]), prprv[,wh,sx,"ap","obs"]*100,
+          type="l", col=gray(0.6), lty=1, lwd=2, lend="butt" )
+    }
+ mtext( "Prevalence of DM (%)", side=2, line=2, las=0, outer=TRUE )
+ mtext( "Age", side=1, line=2, las=0, outer=TRUE )
+   }
> pchk()
```

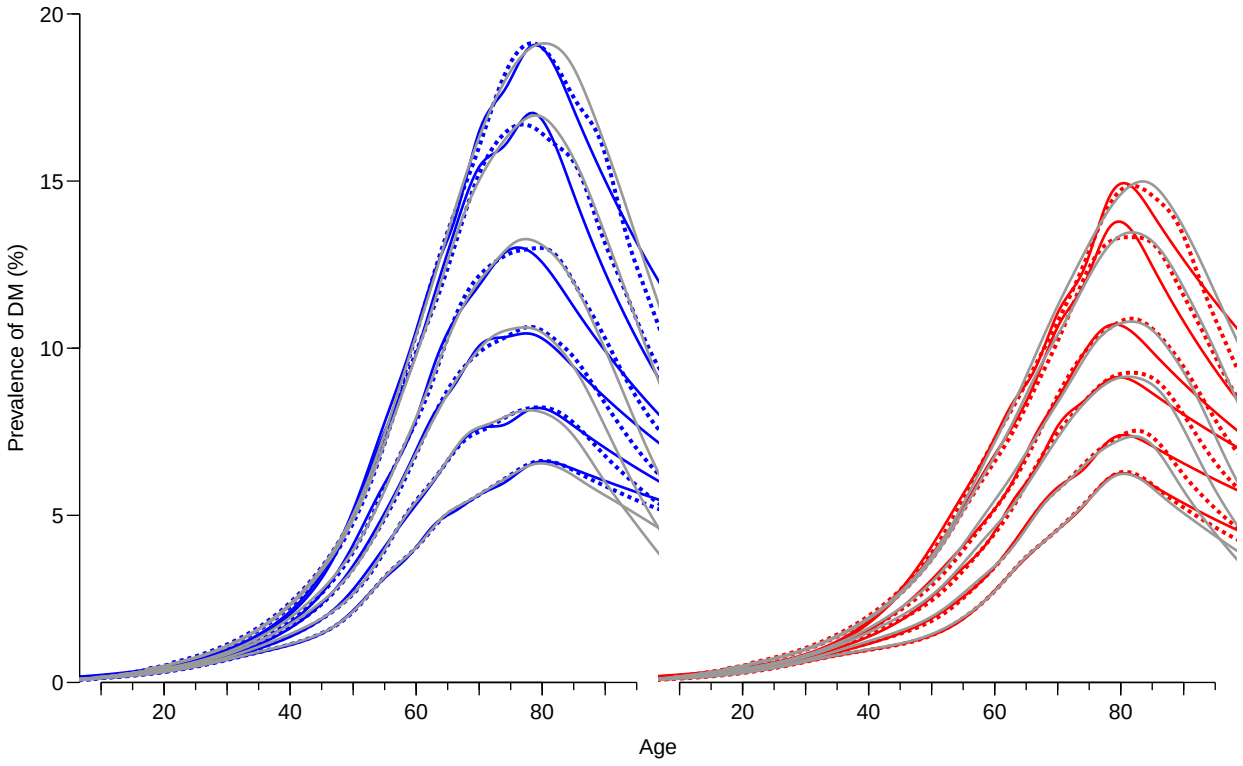


Figure 11.1: *The empirical (smoothed) prevalences at 1 January 1997, 2001,. . . ,2017 (full lines) and the predicted prevalences from using the estimated incidence and mortality rates from an APC-model, computed at 1 month intervals by age and calendar time (broken lines). Gray lines are predictions using an age-perido model for incidence rates. Left panel is men, right panel is women.* `./graph/comp-check`

Since the APC-models for rates provide a better fit (figure **??**), we shall use these in the reporting of the different scenarios.

We now compare the predicted prevalences under the four scenarios at 2016-01-01:

```
> str( prprv )
 num [1:1200, 1:253, 1:2, 1:2, 1:8] 0 0.000541 0.000545 0.000549 0.000553 ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ t   : chr [1:253] "1996" "1996.08333333333" "1996.16666666667" "1996.25" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ mod : chr [1:2] "ap" "apc"
  ..$ what: chr [1:8] "obs" "m-fix" "i-fix" "all-f" ...
> dimnames(prprv)[["t"]][np <- 253]
[1] "2017"
> prprv[floor(dim(prprv)[1]/1.5)+1:5,np,"M","apc",]*100
                  what
a                      obs    m-fix    i-fix    all-f mort inc const org
  66.7083333333333 14.36948 12.83262 11.33567 10.03743   NA  NA    NA  NA
  66.7916666666667 14.41057 12.86292 11.37680 10.06833   NA  NA    NA  NA
  66.875           14.45157 12.89311 11.41782 10.09910   NA  NA    NA  NA
  66.9583333333333 14.49249 12.92319 11.45873 10.12973   NA  NA    NA  NA
  67.0416666666667 14.53333 12.95316 11.49953 10.16022   NA  NA    NA  NA


> par( mfrow=c(1,2), mar=c(0,0,0,0), oma=c(3,4,1,1),
+      mgp=c(3,1,0)/1.6, las=1, bty="n" )
> lpl <- function( sx, cl )
+ matplot( a.pt, cbind(prprv[,np,sx,"apc",c("obs","m-fix","i-fix","all-f")],
+                      prprv[, 1,sx,"apc",1  ])*100,
+          xlim=c(20,90), ylim=c(0,20), yaxs="i", yaxt="n",
+          xlab="Age", ylab="Prevalence (%)",
+          type="l", lty=c("solid","22","33")[c(1,2,1,3,1)],
+          col=cl,   lwd=c(5,5,3,3,1) )
> lpl( "M", "blue" )
> axis( side=2 )
> axis( side=2, at=0:18, labels=NA, tcl=-0.3)
> lpl( "W", "red" )
> dimnames(prprv)[[5]]
[1] "obs"   "m-fix" "i-fix" "all-f" "mort"  "inc"   "const" "org"
```

Thus, in figure **??**, the difference between the thick fill and the thick broken is the effect of declining mortality — the difference between the thick and the thin full lines is the effect of increasing incidence rates. Finally, the difference between the thin broken line and the very thin full line (lowest) is the effect of the imbalance between incidence and mortality rates in 1995 (disequilibrium).

```
> scen <- c("Mort changes, Inc changes",
+           "Mort 1996, Inc changes",
+           "Mort changes, Inc 1996",
+           "Mort 1996, Inc 1996",
+           "Prevalence 1996")
> c.a <- dimnames(prprv)[[1]][floor(dim(prprv)[1]/1.34)]
> n.a <- as.numeric( c.a )
> hts <- c(prprv[c.a,np,"M","apc",1:4],
+          prprv[c.a, 1,"M","apc",1])*100
> cau.exp <-
+ function( wh=1:5, fill=FALSE )
+ {
+ pdf( paste( "./graph/comp-DMpr-", paste(wh,collapse=""), if( fill ) "W",
```

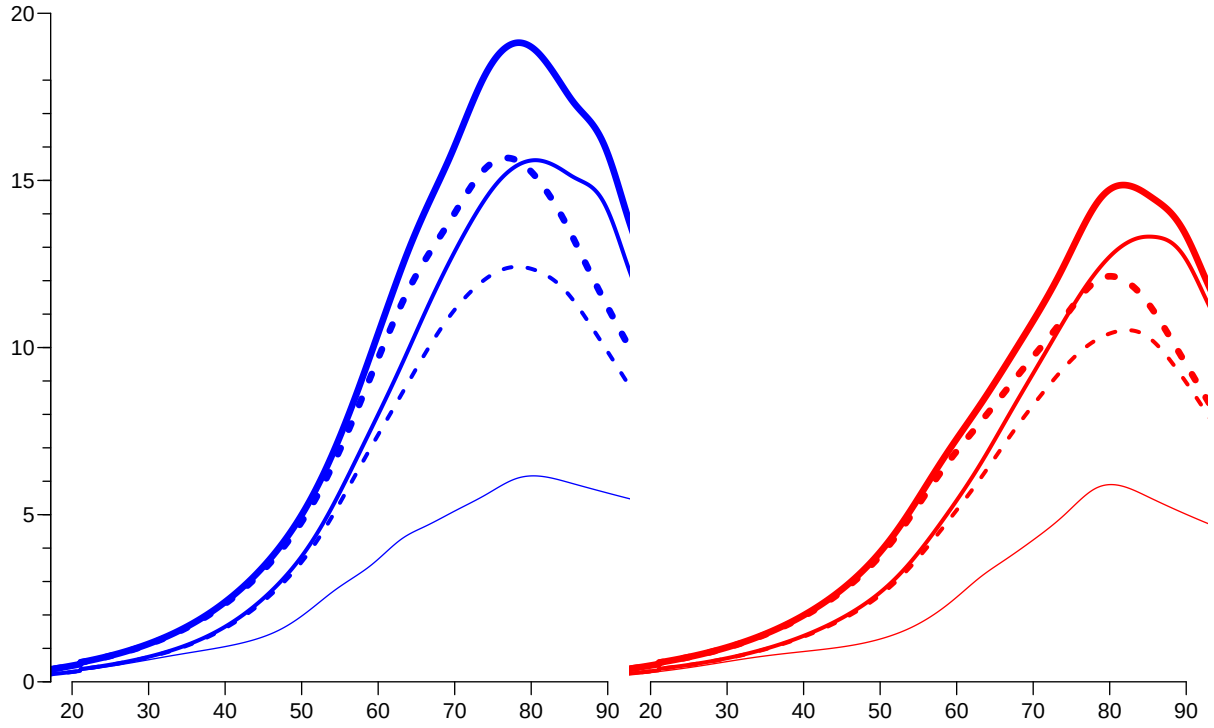


Figure 11.2: *The predicted prevalences under different scenarios:*
Full lines: Mortality rates evolve as observed.
Broken lines: Mortality rates remain as 1996.
Thick lines: Incidence rates evolve as observed.
Thin lines: Incidence rates remain as in 1996.
The very thin lines lowest in the two displays are the observed age-specific prevalences in 1996.
`./graph/comp-causes`

```
+             ".pdf", sep="" ), height=11*9/16, width=11 )
+ par( mfrow=c(1,2), mar=c(0,0,0,0), oma=c(3,4,1,1),
+      mgp=c(3,1,0)/1.6, las=1, bty="n" )
+ matplot( a.pt, cbind(prprv[,np,"M","apc",],prprv[,1,"M","apc",1])*100,
+          xlim=c(20,90), ylim=c(0,20), xlab="Age",
+          ylab="Prevalence (%)", yaxs="i",
+          type="l", lty=rep(c(1,0),2), lwd=c(4,4,2,2,0)+1, col="blue" )
+ axis( side=2, at=0:18, labels=NA, tcl=-0.3 )
+ mtext( "Prevalence (%)", side=2, line=2.5, outer=T, las=0 )
+ matlines( a.pt, prprv[,np,"M","apc",]*100,
+          type="l", lty=rep(c("11","22"),2), lwd=c(4,4,2,2)+1, col="blue" )
+ matlines( a.pt, prprv[,1,"M","apc",]*100,
+           type="l", lty=1, lwd=1, col="blue" )
+ text( rep(20,5)[wh], hts[wh], scen[wh], adj=0, col=gray(0.3), cex=1.4 )
+ for( i in 1:15 )
+ arrows( (20.20+strwidth(scen,cex=1.4))[wh], hts[wh],
+                              rep(n.a,5)[wh], hts[wh],
+         col=gray(0.3), angle=i, lwd=2 )
+ if( fill ) polygon( c(a.pt,rev(a.pt)),
+                     c(prprv[,np,"M","apc",wh[1]],
+                   rev(prprv[,if(wh[2]==5) 1 else np,"M","apc",wh[2]]))*100,
+                     col=rgb(0,0,1,0.3), border="transparent" )
```

```
+ matplot( a.pt, cbind(prprv[,np,"W","apc",],prprv[,1,"W","apc",1])*100,
+          xlim=c(20,90), ylim=c(0,20), xlab="Age", yaxt="n", yaxs="i",
+          type="l", lty=rep(c(1,0),2), lwd=c(4,4,2,2,0)+1, col="red" )
+ matlines( a.pt, prprv[,np,"W","apc",]*100,
+          type="l", lty=rep(c("11","22"),2), lwd=c(4,4,2,2)+1, col="red" )
+ matlines( a.pt, prprv[,1,"W","apc",]*100, type="l", lty=1, lwd=1, col="red" )
+ if( fill ) polygon( c(a.pt,rev(a.pt)),
+                     c(prprv[,np,"W","apc",wh[1]],
+                   rev(prprv[,if(wh[2]==5) 1 else np,"W","apc",wh[2]]))*100,
+                     col=rgb(1,0,0,0.3), border="transparent" )
+ dev.off()
+ }
> cau.exp(1:5)
null device
          1


> for( ff in c(FALSE,TRUE) )
+    {
+ cau.exp(1:2,fill=ff)
+ cau.exp(3:4,fill=ff)
+ cau.exp(4:5,fill=ff)
+ cau.exp(c(1,3),fill=ff)
+ cau.exp(c(2,4),fill=ff)
+    }
```

Figure 11.2 shows the predicted prevalences under 4 different scenarios compared to the observed prevalences as of 1996-01-01.

11.3 How much is attributable to what?

We can compute how much of the age-specific prevalences that are attributable to mortality changes and how much to changes in incidence rates.

The effect of mortality decline can be computed either as the difference between "obs" and "m-fix" or as the difference between "i-fix" and "all-f"; the latter is the mortality effect in a scenario with stable (fixed) incidence rates. There is of course no guarantee that these two quantities are the same.

Similarly the effect of incidence increase can be computed either as the difference between "obs" and "i-fix" or as the difference between "m-fix" and "all-f"; the latter the incidence effect as it would have been if in a scenario with stable (fixed) incidence rates.

Hence we explore how different these quantities are:

```
> par( mfrow=c(1,2), mar=c(3,0,0,0), oma=c(0,4,1,1),
+      mgp=c(3,1,0)/1.6, las=1, bty="n" )
> matplot( a.pt, cbind( prprv[,np,"M","apc","obs"  ]-
+                       prprv[,np,"M","apc","m-fix"],
+                       prprv[,np,"M","apc","i-fix"]-
+                       prprv[,np,"M","apc","all-f"] )*100,
+          xlim=c(20,90), ylim=c(0,6), xlab="Age", ylab="Prevalence (%)",
+          type="l", lty=1, lwd=c(4,2)+1, col="blue", yaxs="i" )
> axis( side=2, at=0:14/2, labels=NA, tcl=-0.3 )
> mtext( "Prevalence difference (%)", side=2, line=2.5, outer=T, las=0 )
> matlines(a.pt, cbind( prprv[,np,"M","apc","obs"  ]-
+                       prprv[,np,"M","apc","i-fix"],
```

```
+                        prprv[,np,"M","apc","m-fix"]-
+                        prprv[,np,"M","apc","all-f"] )*100,
+          type="l", lty="22", lwd=c(4,2)+1, col="blue" )
> matplot( a.pt, cbind( prprv[,np,"W","apc","obs"  ]-
+                        prprv[,np,"W","apc","m-fix"],
+                        prprv[,np,"W","apc","i-fix"]-
+                        prprv[,np,"W","apc","all-f"] )*100,
+          xlim=c(20,90), ylim=c(0,4.5), xlab="Age", yaxt="n", yaxs="i",
+          type="l", lty=1, lwd=c(4,2)+1, col="red" )
> matlines(a.pt, cbind( prprv[,np,"W","apc","obs"  ]-
+                        prprv[,np,"W","apc","i-fix"],
+                        prprv[,np,"W","apc","m-fix"]-
+                        prprv[,np,"W","apc","all-f"] )*100,
+          type="l", lty="22", lwd=c(4,2)+1, col="red" )
```

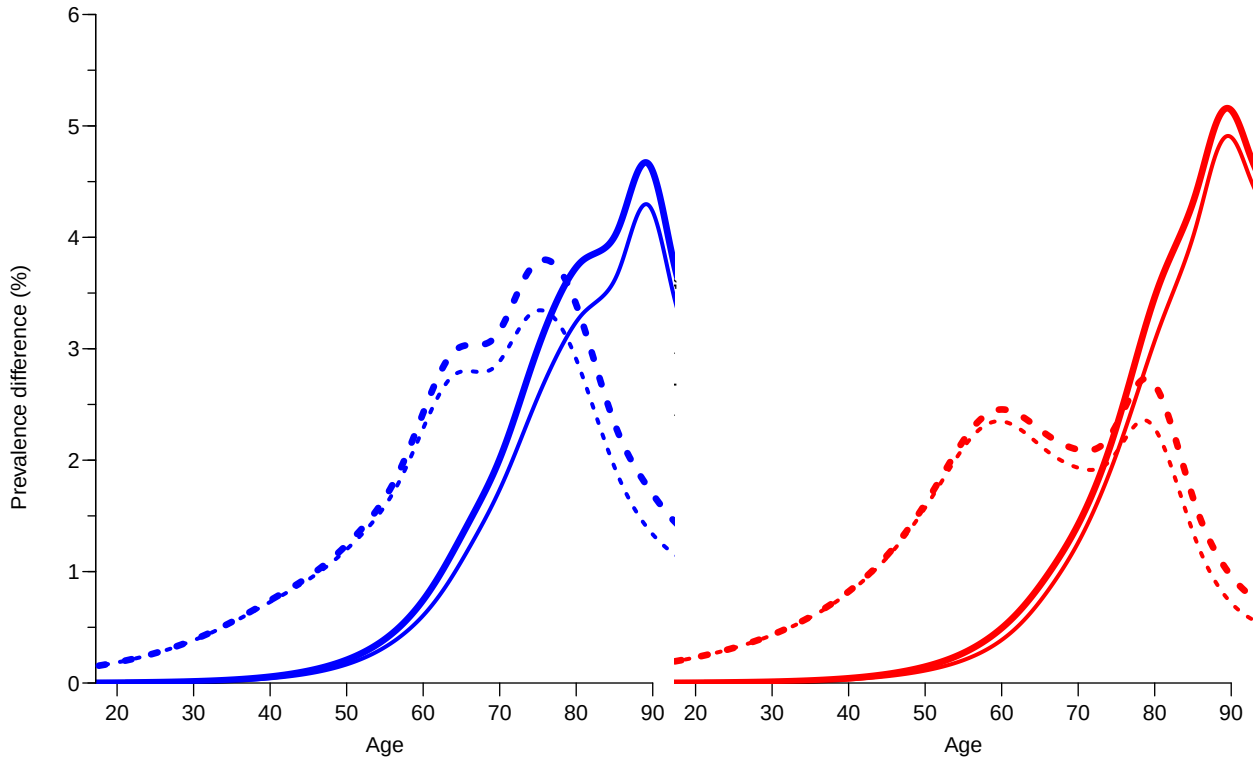


Figure 11.3: *Suggested contributions to age-specific prevalences from increasing mortalities over the period 1996–2016; the thick lines are obtained by subtracting the prediction based on fixing one rate from the one using the observed rates; thin lines based on subtracting the prediction based on fixing both rates from that where only one is fixed. Full lines are for differences attributable to changes in mortality rates, broken lines are for changes attributable to changes in incidence rates.* `./graph/comp-attrib`

From figure 11.3 we see that the two different ways of computing the contribution give pretty much the same results — the differences never exceed 0.5%. Therefore, if we want to attribute fractions of the prevalence in 2016 to decreasing mortality and increasing incidence respectively, we would want two measures that had a sum equal the the difference between the scenario with observed mortality and incidence rates ("`obs`"), and the scenario with rates fixed to those from 1996 ("`all-f`"). This is obtained by taking the average of the two curves in each scenario.

The thin lines at the bottom of figure 11.2 represent the prevalence at 1996-01-01, so it is pretty clear that the incidence an mortality rates as observed by 1996 did not provide for at steady state.

So basically we can subdivide the prevalence at any point in time into 4 components:

1. the "inherited" prevalences from 1996.
2. the prevalence attributable to rates of mortality and incidence as of 1996 — the "epidemiological disequilibrium" as of 1996.
3. the prevalence attributable to the *change* in the incidence rates.
4. the prevalence attributable to the *change* in the mortality rates.

So we now fill out the remaining 3 dimensions of `prprv` according to this:

```
> prprv[,,,,"mort" ] <- ( prprv[,,,,"obs"  ]-prprv[,,,,"m-fix"] +
+                         prprv[,,,,"i-fix"]-prprv[,,,,"all-f"] ) / 2
> prprv[,,,,"inc"  ] <- ( prprv[,,,,"obs"  ]-prprv[,,,,"i-fix"] +
+                         prprv[,,,,"m-fix"]-prprv[,,,,"all-f"] ) / 2
> prprv[,,,,"const"] <-   prprv[,,,,"all-f"]-prprv[,rep(1,dim(prprv)[2]),,,"obs"]
```

The component `obs` of `prprv` is the total prevalence (from the modeling), whereas `const`, `inc` and `mort` are the three components. It would be more logical if the `obs` were the last component:

```
> prprv[,,,,"org"] <-
+ prprv[,,,,"obs"] - ( prprv[,,,,"mort"] +
+                      prprv[,,,,"inc"]  +
+                      prprv[,,,,"const"] )
```

The components `obs`, `const`, `inc` and `mort` now together make up the total prevalence of diabetes (as proportions) for a given combinations of sex, age and date. Thus we can show these for each of the 21 dates 1 January 1996,...,2016.

First we define a function to make the component plots, and then we can plot the resulting development for men and women, for convenience we also put the latter in a function.

```
> poly.parts <-
+ function( x, crv, col, xlim, ylim, txt="" )
+ {
+ crv <- t(apply(cbind(0,crv),1,cumsum))
+ matplot( x, crv, type="n", xaxt="n", yaxt="n", xlab="", ylab="",
+          xlim=xlim, ylim=ylim, yaxs="i", bty="n" )
+ for( i in 2:ncol(crv) )
+ polygon( c(x,rev(x)), c(crv[,i],rev(crv[,i-1])),
+          col=col[i-1], border=col[i-1])
+ text( par("usr")[1:2]%*%c(0.1,0.9),
+       par("usr")[3:4]%*%c(0.9,0.1), txt, adj=c(1,0), font=2 )
+ }
> one.comp <-
+ function( sex, clr )
+ {
+ par( mfrow=c(3,6), mar=c(0,0,0,0), oma=c(3,3,1,1), mgp=c(3,1,0)/1.6 )
+ yn <- round(seq(1,dim(prprv)[2],,21))
+ for(j in 1:21)
+ {
```

```
+ poly.parts( a.pt,
+             prprv[,yn[j],sex,"apc",c("org","const","inc","mort")]*100,
+             col=clr, xlim=c(20,90), ylim=c(0,25),
+             txt=dimnames(prprv)[[2]][yn[j]] )
+ abline(h=0)
+ if( j==1 ) text( rep(25,3), c(13,15,17)+0.5,
+                  c("Imbalance","Incidence","Mortality"),
+                  col=clr[2:4], font=2, adj=0, cex=1.2 )
+ if( j %in% c(1,7,13) ) axis( side=2 )
+ if( j %in% 13:18 ) axis( side=1 )
+ mtext( "Age", side=1, outer=TRUE, line=1.5, font=1, las=0 )
+ mtext( "Prevalence of DM", side=2, outer=TRUE, line=1.5, font=1, las=0 )
+ }
+ }
```

With these functions in place we make a graph of the prevalence of DM by 2016 subdivided by component of change since 1996:

```
> par( mfrow=c(1,2), mar=c(1,0,1,0), mgp=c(3,1,0)/1.6, las=1,
+      oma=c(2,3,0,1), bty="n" )
> clr <- rgb(c(3,2,1.5,0)/3,c(3,2,1.5,0)/3,1)
> poly.parts( a.pt, prprv[,"2017","M","apc",c("org","const","inc","mort")]*100,
+             col=clr, xlim=c(20,90), ylim=c(0,20) )
> abline(h=0:22,v=2:9*10,col=gray(0.9))
> axis( side=1 )
> axis( side=2 )
> axis( side=2, at=1:18, labels=NA, tcl=-0.3 )
> text( rep(25,3), 17:19-1.5,
+       c("Imbalance","Incidence","Mortality"),
+       col=clr[2:4], font=2, adj=0, cex=1.0 )
> clr <- rgb(1,c(3,2,1.5,0)/3,c(3,2,1.5,0)/3)
> poly.parts( a.pt, prprv[,"2017","W","apc",c("org","const","inc","mort")]*100,
+             col=clr, xlim=c(20,90), ylim=c(0,20) )
> abline(h=0:22,v=2:9*10,col=gray(0.9))
> axis( side=1 )
> text( rep(25,3), 17:19-1.5,
+       c("Imbalance","Incidence","Mortality"),
+       col=clr[2:4], font=2, adj=0, cex=1.0 )
> mtext( "Age", side=1, outer=TRUE, line=0.8, font=1, las=0 )
> mtext( "Prevalence of DM (%)", side=2, outer=TRUE, line=2, font=1, las=0 )
```

11.4 Number of diabetes patients in Denmark

In the previous section we only looked at the age-specific prevalences, because these are the quantities that are driven by the incidence and mortality rates. However, it is also of interest to see how the actual number of diabetes patients would have looked under the different scenarios, specifically how the *number* of the current patients that can be attributed to the various components.

Also note that since the previous calculations were for age-specific prevalences we have a constant reference as the prevalences at 1996, but when we multiply by the population figures we would of course see differences in numbers and age-distribution of the diabetes population even if the age-specific prevalences were unchanged.

To show these effects we set up an array `prn` with `structure` (but not extent) as `prprv` to hold the number of diabetes patients by category, assuming the age-distribution in the

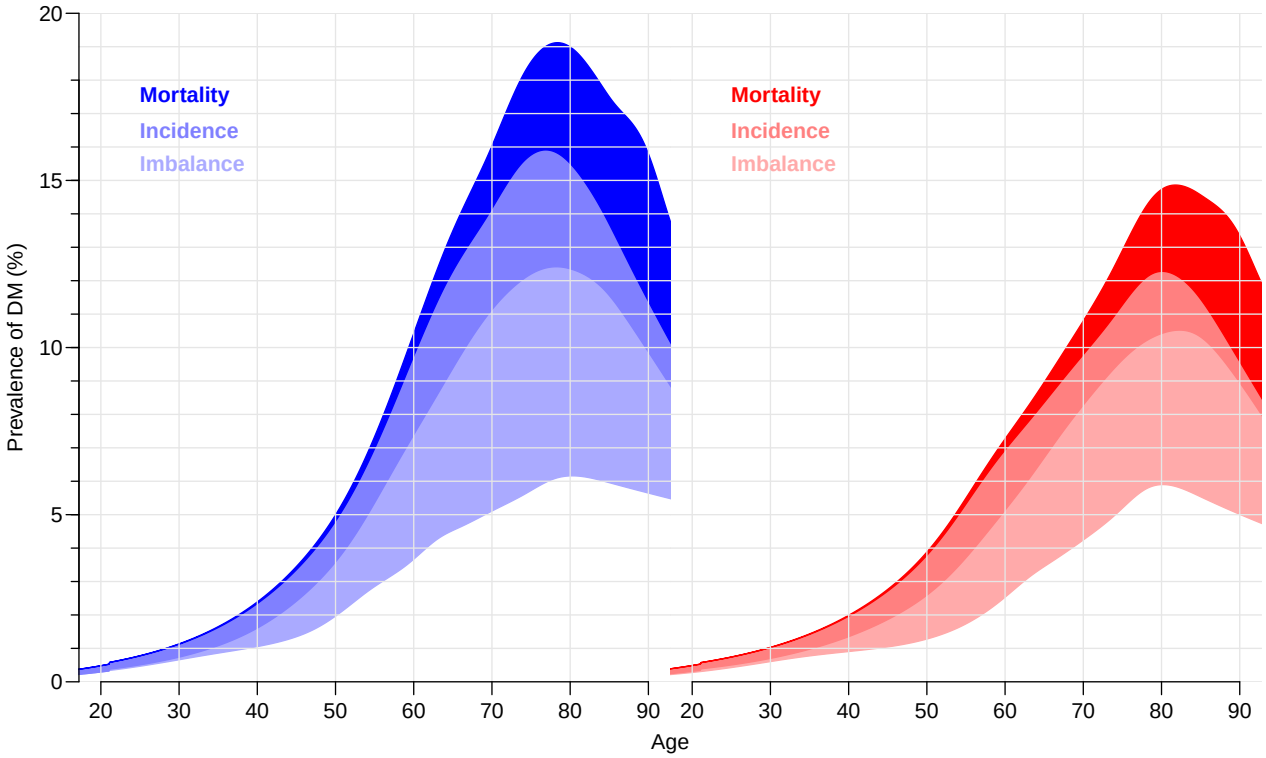


Figure 11.4: *Changes in predicted age-specific prevalences of DM in Denmark 2016 among men (blue) and women (red), partitioned by the contribution from incidence and mortality decrease, in since 1996 and from disequilibrium between incidence and mortality at 1996 ("Imbalance").* `./graph/comp-prev-comp-2016`

population to be as actually observed (that is as extracted from Statistics Denmark). However `prn` will have 100 age-classes rather than 1200 (100/`int`), and only 21 dates: `prprv`.

```
> dn <- dimnames(prprv)
> dn[[1]] <- 0:99
> dn[[2]] <- 1996:2017
> dn[[5]] <- dn[[5]][5:8]
> prn <- NArray( dn )
> str(prprv)
 num [1:1200, 1:253, 1:2, 1:2, 1:8] 0 0.000541 0.000545 0.000549 0.000553 ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ t   : chr [1:253] "1996" "1996.08333333333" "1996.16666666667" "1996.25" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ mod : chr [1:2] "ap" "apc"
  ..$ what: chr [1:8] "obs" "m-fix" "i-fix" "all-f" ...

> str(prn)

 logi [1:100, 1:22, 1:2, 1:2, 1:4] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:100] "0" "1" "2" "3" ...
  ..$ t   : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ mod : chr [1:2] "ap" "apc"
  ..$ what: chr [1:4] "mort" "inc" "const" "org"
```

In order to fill in the numbers we use the estimated age-specific prevalences at 1st January each year, that is at the dates 1996-01-01,...,2016-01-01 in the entries along the `t`-dimension of `prprv`. Moreover, we want the prevalences for 1 year age class rather than age-classes of length `int`. So we take the average prevalences from `prprv` over each one-year age-interval. The vectors `wh.a` and `wh.p` will hold the number of the age and period classes from `prprv` which have the desired prevalences (as proportions) that we will use for multiplication with the population figures:

```
> prprv.p <- as.numeric( dimnames(prprv)[["t"]] )
> prn.p <- as.numeric( dimnames(prn)[["t"]] )
> ( wh.p <- match( prn.p, prprv.p ) )
 [1]   1  13  25  37  49  61  73  85  97 109 121 133 145 157 169 181 193 205 217 229 241 253
> prprv <- pmax( prprv, 0 )
```

Now `wh.p` contains the numbers on the date dimension in `prprv.p` that we shall use. But we want the prevalences for 1 year age classes rather than for age-classes of length `int`. So we take the average prevalences from `prprv.p` over each one-year age-interval.

This is really just a simple matrix operation; take a diagonal matrix of 1/12 (well, `int`), and repeat each column 12 (1/`int`) times:

```
> dd <- diag(100)[,rep(1:100,each=1/int)]*int
> dim(dd)
[1]  100 1200
> dd[1:3,c(1,2,12:13,24:25)]
           [,1]       [,2]       [,3]       [,4]       [,5]       [,6]
[1,] 0.08333333 0.08333333 0.08333333 0.00000000 0.00000000 0.00000000
[2,] 0.00000000 0.00000000 0.00000000 0.08333333 0.08333333 0.00000000
[3,] 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.08333333
```

Pre-multiplying this 100×1200 matrix to the 1200 ($= 100/$`int`) age-specific prevalences gives the average prevalences in the 100 1-year age-classes. So we just select the dates at which we want the prevalences:

```
> for( sx in dimnames(prn)[[3]] )
+ for( im in dimnames(prn)[[4]] )
+ for( mm in dimnames(prn)[[5]] )
+ prn[,,sx,im,mm] <- dd %*% prprv[,wh.p,sx,im,mm]
```

Now `prn` contains the estimated prevalences (as fractions) for 100 age classes and the 21 dates. We need to multiply these prevalences by the population figures at these times. The population figures are in `pop`

```
> load( "../nydata/prevN.Rda" )
> str( prp <- xtabs( N ~ A + P + sex, data = prN ) )
 'xtabs' num [1:100, 1:22, 1:2] 36271 36169 35124 35398 33895 ...
 - attr(*, "dimnames")=List of 3
  ..$ A  : chr [1:100] "0" "1" "2" "3" ...
  ..$ P  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ sex: chr [1:2] "M" "W"
 - attr(*, "call")= language xtabs(formula = N ~ A + P + sex, data = prN)
> fCtable( addmargins( xtabs( N ~ P + sex, data = prN ), 2),  w=9 )
```

```
     sex         M         W       Sum
P
1996     2,603,806 2,667,095 5,270,901
1997     2,615,773 2,677,878 5,293,651
1998     2,625,987 2,686,520 5,312,507
1999     2,635,431 2,695,107 5,330,538
2000     2,643,898 2,702,503 5,346,401
2001     2,654,573 2,711,860 5,366,433
2002     2,664,792 2,720,979 5,385,771
2003     2,673,711 2,728,072 5,401,783
2004     2,681,170 2,734,339 5,415,509
2005     2,688,755 2,741,288 5,430,043
2006     2,697,410 2,748,777 5,446,187
2007     2,708,086 2,757,302 5,465,388
2008     2,722,434 2,768,781 5,491,215
2009     2,740,248 2,784,034 5,524,282
2010     2,751,770 2,796,211 5,547,981
2011     2,764,647 2,808,297 5,572,944
2012     2,776,767 2,818,978 5,595,745
2013     2,789,506 2,829,382 5,618,888
2014     2,804,242 2,841,571 5,645,813
2015     2,825,552 2,857,550 5,683,102
2016     2,853,820 2,878,890 5,732,710
2017     2,880,935 2,901,770 5,782,705
```

```
> str( prp )
```

```
 'xtabs' num [1:100, 1:22, 1:2] 36271 36169 35124 35398 33895 ...
 - attr(*, "dimnames")=List of 3
  ..$ A  : chr [1:100] "0" "1" "2" "3" ...
  ..$ P  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ sex: chr [1:2] "M" "W"
 - attr(*, "call")= language xtabs(formula = N ~ A + P + sex, data = prN)
```

```
> str( prn[,,,1,1] )
```

```
 num [1:100, 1:22, 1:2] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 3
  ..$ a  : chr [1:100] "0" "1" "2" "3" ...
  ..$ t  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ sex: chr [1:2] "M" "W"
```

Hence we multiply the population figures by the prevalences to get the total number of persons in the different groups:

```
> for( i in dimnames(prn)[[4]] )
+ for( j in dimnames(prn)[[5]] )
+ prn[,,,i,j] <- prn[,,,i,j] * prp
> dmp <- apply( prn[,,,"apc",], 1:3, sum )
> str( dmp )
```

```
 num [1:100, 1:22, 1:2] 18.6 22.1 23.4 25.8 26.9 ...
 - attr(*, "dimnames")=List of 3
  ..$ a  : chr [1:100] "0" "1" "2" "3" ...
  ..$ t  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ sex: chr [1:2] "M" "W"
```

First we draw a simple pyramid of the age-distribution of diabetes patients in Denmark:

```
> # Note: This uses the undocumented feature that if the first
> # number in a column is negative this is taken as the left endpoint of
> # the bar. So c(-m,m,f) is a bar starting at -m, and a division at
> # -m+m(=0) and an upper end at -m+m-f. Coloring is from the top, that is
> # the part stretching from -m+m to -m+m+f get the first color
> pp <- "2016"
> oo <- c("mort","inc","const","org")
> lim <- 6
> clr <- c("red","blue")
> draw.dmp <-
+ function(pp)
+ {
+ par( mar=c(3,3,3,0), mgp=c(3,1,0)/1.6, las=1 )
+ barplot( height=t( cbind( -dmp[,pp,"M"],
+                           dmp[,pp,"M"],
+                           dmp[,pp,"W"] ) )/ 1000,
+          horiz=TRUE, col=clr,
+          border="transparent",space=0,axes=FALSE,
+          names.arg=rep("",dim(prn)[1]),
+          xlim=c(-1,1)*lim*1.05,
+          xlab="Persons in 1 year class (1000s)",ylab="Age")
+ abline(h=seq(0,100,5),
+        v=seq(-lim,lim,0.5),
+        col="white")
+ axis( side=1, at=seq(-lim,lim,1), labels=abs(seq(-lim,lim,1)) )
+ axis( side=2, at=seq(0,100,20) )
+ axis( side=2, at=seq(0,100,5), labels=NA, tcl=-0.3 )
+ mtext( pp, at=-lim, adj=1.4, cex=1.3, font=1 )
+ mtext( formatC(sum(dmp[,pp,"M"]),0,format="f",big.mark=","),
+        at=-1, col="blue", line=0, cex=0.99 )
+ mtext( formatC(sum(dmp[,pp,"W"]),0,format="f",big.mark=","),
+        at= 1, col="red" , line=0, cex=0.99 )
+ mtext( "N", at=0, line=0, cex=0.99 )
+ }
> pdf( "./graph/comp-obs-film.pdf", width=8, height=6 )
> for( pp in paste(1996:2017) ) draw.dmp(pp)
> dev.off()
null device
          1


> for( pp in paste(seq(1996,2017,3)) )
+ {
+ pdf( paste("./graph/comp-obs-", pp, ".pdf", sep=""), width=8, height=6 )
+ draw.dmp(pp)
+ dev.off()
+ }
```

Now we can also draw a population pyramid of the *predicted* number of DM patients using colors that range from very light to full:

```
> shd <- c(0.0, 1.1, 2.0, 2.8) / 3
> een <- rep(1,4)
> clr <- rgb( c(een,rev(shd)),
+             c(shd,rev(shd)),
+             c(shd,    een ) )
> clr
```

```
[1] "#FF0000" "#FF5E5E" "#FFAAAA" "#FFEEEE" "#EEEEFF" "#AAAAFF" "#5E5EFF" "#0000FF"
> oo <- c("mort","inc","const","org")
> lim <- 6
> draw.pyr <-
+ function(pp)
+ {
+ par( mar=c(3,3,3,0), mgp=c(3,1,0)/1.6, las=1 )
+ barplot( height=t( cbind( -apply(prn[,pp,"M","apc",       ], 1, sum),
+                                 prn[,pp,"M","apc",    oo ],
+                                 prn[,pp,"W","apc",rev(oo)] ) )/ 1000,
+          horiz=TRUE, col=clr[c(1,8:2)], border=rep("transparent",8),
+          space=0, axes=FALSE, names.arg=rep("",dim(prn)[1]),
+          xlim=c(-1,1)*lim*1.05,
+          xlab="Persons in 1 year class (1000s)",ylab="Age")
+ abline(h=seq(0,100,5),
+        v=seq(-lim,lim,0.5),
+        col="white")
+ axis( side=1, at=seq(-lim,lim,1), labels=abs(seq(-lim,lim,1)) )
+ axis( side=2, at=seq(0,100,20) )
+ axis( side=2, at=seq(0,100,5), labels=NA, tcl=-0.3 )
+ tt <- addmargins( apply( prn[,pp,,"apc",],2:3, sum ), 2 )
+ nn <- tt / tt[,5] * 100
+ ppos <- seq(1,5.9,,5)-0.1
+ npos <- -rev(ppos)
+ mtext( pp, at=-lim, adj=1.8, line=2, cex=1.2, font=1 )
+ mtext( c(lg<- c("Mort","Inc","Imbal","Org","All"),rev(lg)),
+        at=c(npos,ppos), col="black", cex=0.99, line=2 )
+ mtext( formatC(tt["M",1:5],0,,"f",,,","),
+        at=npos, col="blue", line=1, cex=0.99 )
+ mtext( formatC(tt["W",5:1],0,,"f",,,","),
+        at=ppos, col="red" , line=1, cex=0.99 )
+ mtext( formatC(nn["M",1:4],1,4,"f"),
+        at=npos[1:4], col="blue", line=0, cex=0.99 )
+ mtext( formatC(nn["W",4:1],1,4,"f"),
+        at=ppos[2:5], col="red" , line=0, cex=0.99 )
+ mtext( "N", at=0, line=1, cex=0.99 )
+ mtext( "%", at=0, line=0, cex=0.99 )
+ }
> pdf( "./graph/comp-DMpr-film.pdf", width=9, height=6 )
> for( pp in paste(1996:2017) ) draw.pyr(pp)
> dev.off()
null device
          1

> for( pp in paste(seq(1996,2017,3)) )
+ {
+ pdf( paste("./graph/comp-DMpr-", pp, ".pdf", sep=""), width=8, height=6 )
+ draw.pyr(pp)
+ dev.off()
+ }
```

11.5 Time trends in the components

It is of course also of interest to see how large a fraction of the DM-patients in various ages that can be attributed to the different components at different times.

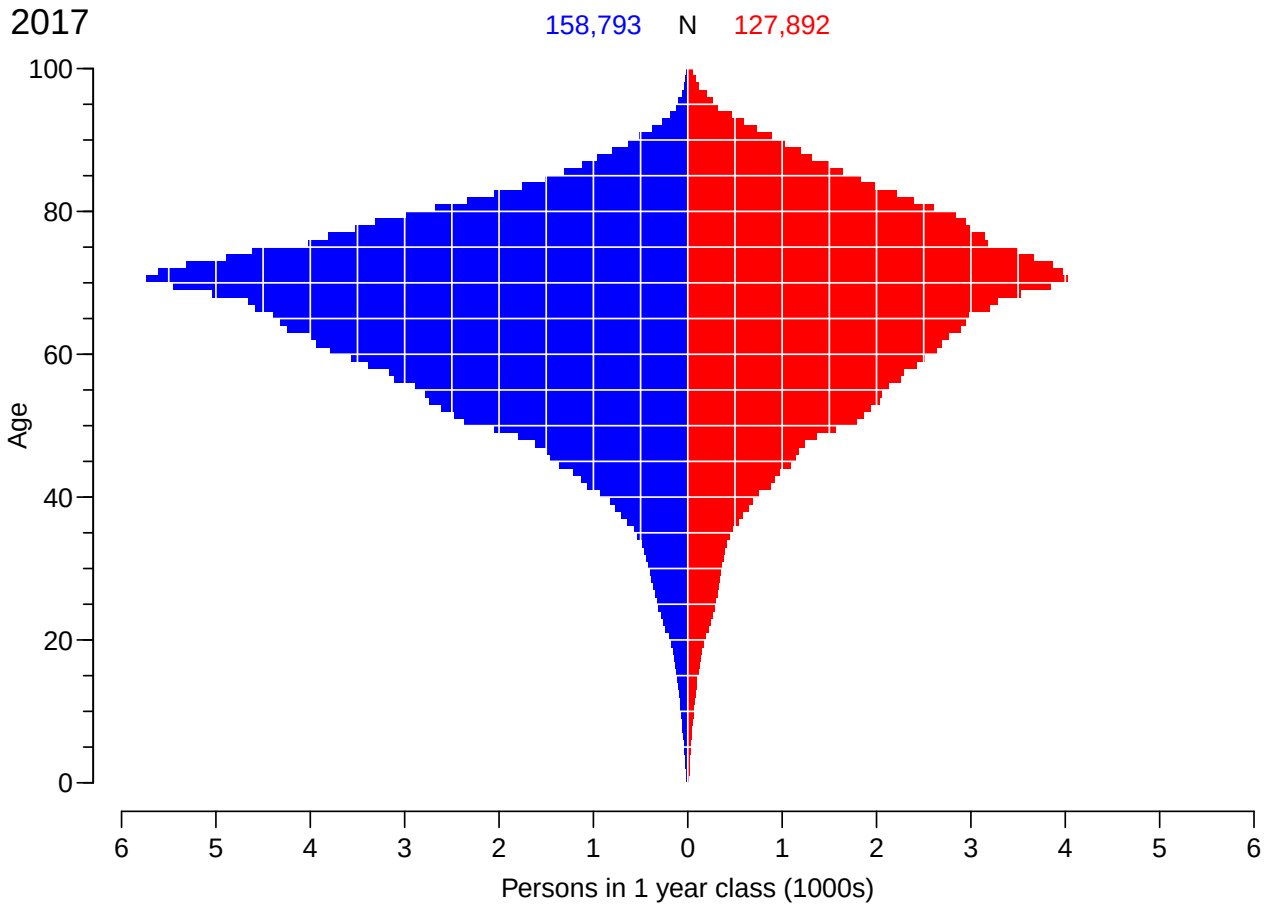


Figure 11.5: *Empirical age-distribution of the diabetes cases in Denmark as of 2017-01-01.*
`./graph/comp-obs-2017`

```
> str( prprv )
 num [1:1200, 1:253, 1:2, 1:2, 1:8] 0 0.000541 0.000545 0.000549 0.000553 ...
 - attr(*, "dimnames")=List of 5
  ..$ a   : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ t   : chr [1:253] "1996" "1996.08333333333" "1996.16666666667" "1996.25" ...
  ..$ sex : chr [1:2] "M" "W"
  ..$ mod : chr [1:2] "ap" "apc"
  ..$ what: chr [1:8] "obs" "m-fix" "i-fix" "all-f" ...
> dimnames( prprv )[[5]]
[1] "obs"   "m-fix" "i-fix" "all-f" "mort"  "inc"   "const" "org"
```

We extract the prevalences at ages 60, 70 and 80, and compute the fractions of all DM patients attributable to each component:

```
> aloc <- match( 6:8*10, floor(as.numeric(dimnames(prprv)[[1]])) )
> ptrend <- ( prprv[aloc,,,"apc",-(2:4)] + prprv[aloc-1,,,"apc",-(2:4)] )/2
> str( ptrend )
 num [1:3, 1:253, 1:2, 1:5] 0.0367 0.0511 0.0616 0.037 0.0515 ...
 - attr(*, "dimnames")=List of 4
  ..$ a   : chr [1:3] "60.0416666666667" "70.0416666666667" "80.0416666666667"
  ..$ t   : chr [1:253] "1996" "1996.08333333333" "1996.16666666667" "1996.25" ...
```

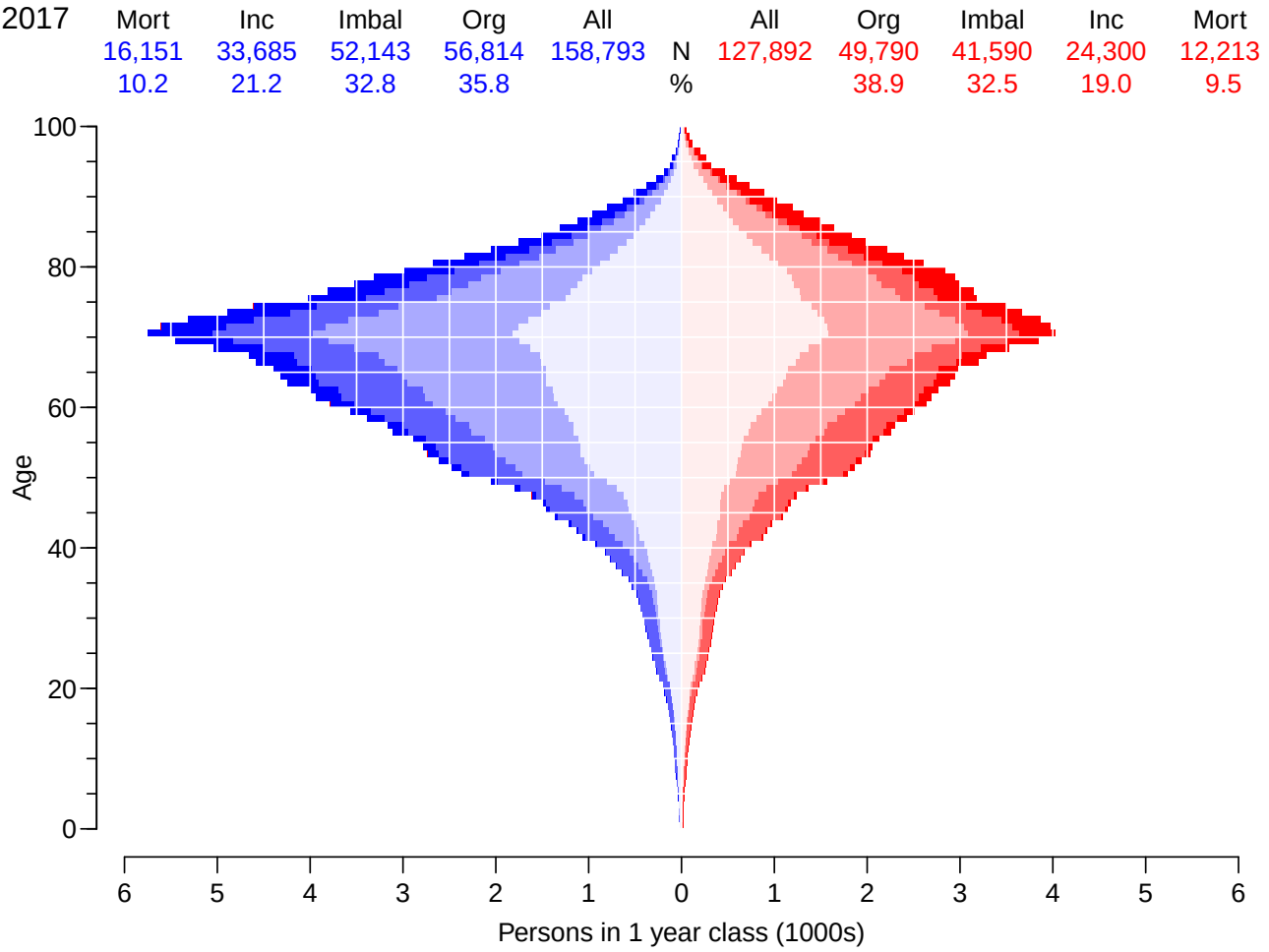


Figure 11.6: *Age-distribution of the predicted no. of diabetes cases in Denmark as of 2016-01-01, subdivided by the components of disease prevalence:*
Mort: decrease in mortality,
Inc: increase in incidence,
Imbal: constant rates from 1996 (non-steady-state imbalance between incidence and mortality in 1996
Org: age-specific prevalences at 1996-01-01. `./graph/comp-DMpr-2017`

```
  ..$ sex : chr [1:2] "M" "W"
  ..$ what: chr [1:5] "obs" "mort" "inc" "const" ...
```

```
> # Fraction of all DM at each age
> ptrend[,,,-1] <- ptrend[,,,-1]/ptrend[,,,rep("obs",4)]
```

We can now plot the stacked fractions of the components for ages 60, 70 and 80 separately for men and women:

```
> p.pt <- as.numeric(dimnames(ptrend)[[2]])
> par( mfcol=c(3,2), mar=c(0,0,0,2), oma=c(3,4,1,0), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> for( sx in c("M","W") )
+ for( ag in 1:3 )
+ {
+ plot( NA, xlim=c(1996,2017), ylim=0:1*100,
```

```
+           xaxs="i", yaxs="i", xaxt="n", yaxt="n",xlab="",ylab="" )
+ #if( sx=="M" ){
+ #axis( side=2 )
+ #axis( side=2, at=1:9*10, labels=NA ) }
+ if( ag==3 ){
+ axis( side=1 )
+ axis( side=1, at=1:21+1996, labels=NA, tcl=-0.3 ) }
+ polygon( c(p.pt,rev(p.pt)), c( ptrend[ag,,sx,"mort"] *100,
+                              rev(ptrend[ag,,sx,"mort"])*  0),
+          col=clr[if(sx=="M") 8 else 1], border="transparent" )
+ polygon( c(p.pt,rev(p.pt)), c(ptrend[ag,,sx,"mort"],
+                             rev(ptrend[ag,,sx,"mort"]+
+                                 ptrend[ag,,sx,"inc" ]))*100,
+          col=clr[if(sx=="M") 7 else 2], border="transparent" )
+ polygon( c(p.pt,rev(p.pt)), c(ptrend[ag,,sx,"mort"]+
+                                 ptrend[ag,,sx,"inc" ],
+                             rev(ptrend[ag,,sx,"mort"]+
+                                 ptrend[ag,,sx,"inc" ]+
+                                 ptrend[ag,,sx,"const"]))*100,
+          col=clr[if(sx=="M") 6 else 3], border="transparent" )
+ abline( v=seq(2000,2015,5), h=1:9*10, col="white" )
+ axis( side=4, at=1:7*10, tcl=-0.5 )
+ axis( side=4, at=0:14*5, tcl=-0.4, labels=NA )
+ axis( side=4, at=0:70  , tcl=-0.2, labels=NA )
+ text( 1997, 75, paste( if(sx=="W") "Women\n" else "Men\n","age",(6:8*10)[ag]),
+       cex=1.2, font=2, adj=0 )
+ box(col=gray(0.7),bty="o")
+ }
> mtext( side=1, "Date", line=2, outer=TRUE, cex=0.67 )
> mtext( side=2, "Prevalence component (%)", line=2, outer=TRUE, cex=0.67, las=0 )


> elapsed()
----------------------
2019-01-06 at 12:37:13
Time elapsed: 00:00:41
----------------------
```

. . . now input from `yll-th.tex`

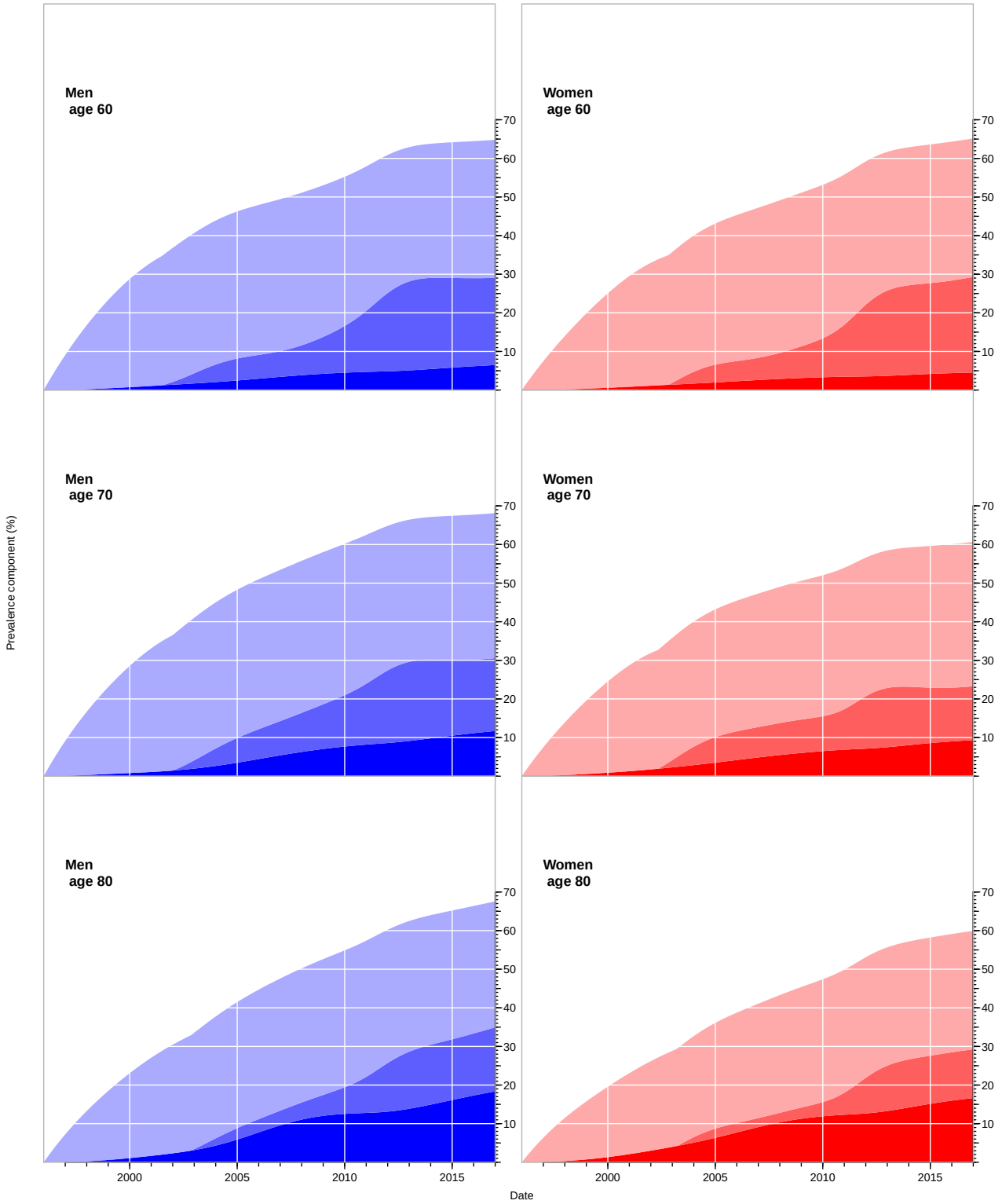


Figure 11.7: *Fraction of the prevalent cases at different times attributable to a) declining mortality (bottom, full color), b) increasing incidence (middle, pale color) and c) prevalence/mortality imbalance at 1996 (top, weak color). The white area above the curves corresponds to the fraction of the cases that would have been around if incidence and mortality rates had remained as in 1996.* `./graph/comp-frac`

Chapter 12

Lifetime lost to diabetes

12.1 Theory

12.1.1 Overall mortality (any cause)

The years of life lost to a given condition is normally defined as the difference between the expected residual lifetime between a person without the condition and a person with the condition.

This amounts to the difference in areas under two (conditional) survival curves — clearly an index age is needed.

Normally this is computed using only age as the time scale and the years of life lost is computed at a given index ages — do for example the years of life lost to diabetes for a 50 year old man is computed.

So what is needed is the survival curves for persons with and without diabetes. For a person with diabetes it is pretty straightforward to compute the survival function conditional on survival to a particular age — it is a simple function of the overall mortality rates: $S(a) = \exp(-\int_0^a \lambda(s)\,\mathrm{d}s)$

But for a person at a given age without diabetes we could use one of the following:

- the naive survival function based on the mortality rates for persons without diabetes — the so-called *net* survival.

  Note that this is *not* what would be the survival function of persons immune to the disease (which in practice would involve not only conditioning on the future — absence of diabetes diagnosis, but also on the unobservable future — absence of diagnosis of diabetes after death)

- the survival based on the total survival for the population, corresponding to the expected survival for a randomly selected persons from the *entire* population

- the multistate based survival for a person without diabetes (at a given age), but taking into account that the person may contract diabetes later and thus incur a higher mortality

12.1.2 Years of life lost by cause of death

The years of life lost is the difference between the areas *below* the survival curves, but of course also (minus) the difference of the areas *above* the survival curves.

Now, if several causes of death are recorded, the area over the survival curve can be subdivided by cause of death. The lifetime lost by cause of death is using the fact that the difference between the survival probabilities is the same as the difference between the death probabilities. If several causes of death (3, say) are considered then:

$$\begin{aligned} S(a) = 1 &- \mathrm{P}\{\text{dead from cause 1 at } a\} \\ &- \mathrm{P}\{\text{dead from cause 2 at } a\} \\ &- \mathrm{P}\{\text{dead from cause 3 at } a\} \end{aligned}$$

and hence:

$$\begin{aligned} S_{\mathrm{Well}}(a) - S_{\mathrm{Diseased}}(a) = &\mathrm{P}\{\text{dead from cause 1 at } a|\text{Diseased}\} \\ &+ \mathrm{P}\{\text{dead from cause 2 at } a|\text{Diseased}\} \\ &+ \mathrm{P}\{\text{dead from cause 3 at } a|\text{Diseased}\} \\ &- \mathrm{P}\{\text{dead from cause 1 at } a|\text{Well}\} \\ &- \mathrm{P}\{\text{dead from cause 2 at } a|\text{Well}\} \\ &- \mathrm{P}\{\text{dead from cause 3 at } a|\text{Well}\} \end{aligned}$$

So we can conveniently define the lifetime lost from age a due to cause 2, say, by:

$$\begin{aligned} \mathrm{LL}_2(a) = \int_a^\infty &\mathrm{P}\{\text{dead from cause 2 at } u|\text{Diseased \& alive at } a\} \\ &- \mathrm{P}\{\text{dead from cause 2 at } u|\text{Well \& alive at } a\}\, \mathrm{d}u \end{aligned}$$

These quantities have the property that their sum is the total years of life lost due to the disease:

$$\mathrm{LL}(a) = \mathrm{LL}_1(a) + \mathrm{LL}_2(a) + \mathrm{LL}_3(a)$$

The terms inside the integral are computed as

$$\begin{aligned} \mathrm{P}\{\text{dead from cause 2 at } u|\text{Diseased \& alive at } a\} &= \int_a^u \lambda_{2,\mathrm{Dis}}(s) S_{\mathrm{Dis}}(s)/S_{\mathrm{Dis}}(a)\, \mathrm{d}s \\ \mathrm{P}\{\text{dead from cause 2 at } u|\text{Well \& alive at } a\} &= \int_a^u \lambda_{2,\mathrm{Well}}(s) S_{\mathrm{Well}}(s)/S_{\mathrm{Well}}(a)\, \mathrm{d}s \end{aligned}$$

So basically we want to compute the "time spent in death states", but referring to the above discussion, also taking the possibility of disease into account.

12.1.3 Multistate set-up for register data

We have not only death subdivided by cause (CVD, Cancer, Respiratory, Other), but also we have two different — mutually exclusive — disease states, namely T1D and T2D. Thus we have a multistate set-up, so we will use a comprehensive model to compute the years of life

lost to T1D and T2D in terms of different causes of death and also explore how this depends on sex and calendar time.

We can illustrate the multistate model in figure 12.1; we see there are 7 states (3 transient and 4 absorbing) and 14 transitions, 12 of which are to absorbing states.

```
> library(Epi)
> TT <- matrix(NA,7,7)
> colnames(TT) <-
+ rownames(TT) <- c("no DM","T1D","T2D","D-CVD","D-Can","D-Res","D-Oth")
> TT[1,-1] <- 1
> TT[2:3,4:7] <- 1
> TT
      no DM T1D T2D D-CVD D-Can D-Res D-Oth
no DM    NA   1   1     1     1     1     1
T1D      NA  NA  NA     1     1     1     1
T2D      NA  NA  NA     1     1     1     1
D-CVD    NA  NA  NA    NA    NA    NA    NA
D-Can    NA  NA  NA    NA    NA    NA    NA
D-Res    NA  NA  NA    NA    NA    NA    NA
D-Oth    NA  NA  NA    NA    NA    NA    NA
> boxes.matrix( TT, boxpos=list( x=c(20,20,20,rep(80    ,4)),
+                                y=c(50,20,80,seq(90,10,,4)) ), hmult=3 )
```

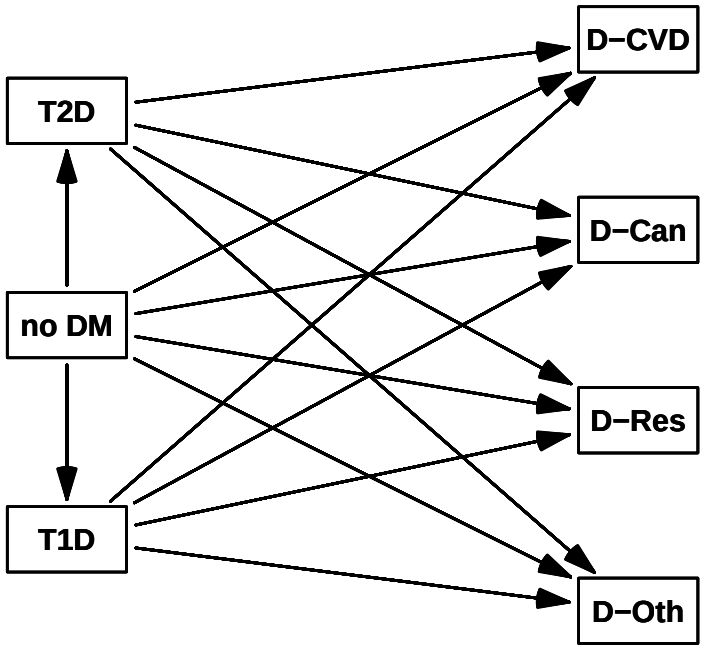


Figure 12.1: *Multistate setup for the lifetime lost analyses. The multistate model is used to compute the conditional survival from select ages starting in each of the three transient stages, noDM, T1D and T2D.* ./graph/yll-boxes

Thus we need to have age-specific rates for each of these transitions; incidence of T1D resp. T2D, and cause-specific mortality rates for persons without DM as well as persons with T1D and T2D, a total of 14 set of rates.

We model the transition rates separately for the two sexes using age-period-cohort models with smooth effects of the three terms.

We want predicted rates for ages in 1-month intervals in order to simplify the calculations of of state occupancy probabilities which are eventually the quantities we shall integrate to get at the lifetime lost to different diseases distributed by cause of death.

12.1.4 Practical implementation of multistate model

The exercise is done using arrays to hold rates, transition probabilities, state probabilities, sojourn times and years of life lost:

`Tr` — predicted transition rates (or more precisely one-month transition probabilities) classified by sex, date, age, from state and to state. Each slice classified by the two last dimensions is a 7 by 7 matrix with values only in the 14 entries corresponding to possible transitions. Moreover the diagonal is filled so that the row-sums of the slice is 1.

`Pr` — state occupancy probabilities classified by sex, date, entry state, entry age, current age and current state. The sum over the last dimension is 1 for any combination of the other dimensions — a person is in one of the states with probability 1.

The array is filled by starting at a given sex, date, entry state and entry age, at the current age equal to entry age and state occupancy equal to 1 for entry state and 0 for all other states, and then updating the state occupancy at subsequent ages by multiplying the state occupancy vector with the relevant slice of the transition array `Tr`.

`Sj` — sojourn times (time spent) in each of the 4 absorbing state from age at entry till infinity (which in practice is 100 years). These are computed as the integral of the state occupancy probabilities in the `Pr` array, integrating from current age equal to age at entry and up to infinity. Computationally we are integrating all the way from 0, but the probabilities are 0 for the absorbing states for age less than entry age.

`YLL` — life time lost (by cause of death) id the difference in sojourn times in the absorbing states between those who start in states T1 resp. T2, and those who start in state noDM. This array is classified by sex, type of diabetes, age and date of entry.

`YLL1` — life time lost during the first year after entry was computed similarly, however only integration over the first year.

`pr` number of prevalent cases of diabetes classified by sex, age (1 month intervals), date (1 January each year) and type of diabetes.

`ic` number of incident cases of diabetes classified by sex, age (1 month intervals), period (1 year) and type of diabetes.

`YLLpr` future life time lost among prevalent cases of diabetes, classified by sex, date (1 January each year), type of diabetes and cause of death. Constructed by multiplying `pr` with `YLL` for each of the absorbing states.

`YLLic` future life time lost among incident cases of diabetes, classified by sex, period (1 year), type of diabetes and cause of death. Constructed by multiplying `ic` with `YLL` for each of the absorbing states.

The burden of diabetes in terms of life time lost is shown by graphing relevant slices of the arrays.

. . . now input from `yll-cod.tex`

12.2 Models for rates

First we read the base-dataset with complete follow-up for the entire Danish population — in this analysis we omit duration of diabetes, because we do not have reliable estimates of duration effects beyond 20 years

```
> library( Epi )
> start()
---------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-05-27 15:42:01
---------------------------------------
> load( file="../nydata/rt.Rda" )
> load( file="../nydata/inits.Rda" )
> with( rt, summary(Dcvd+Dcan+Dres+Doth-D) )
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
      0       0       0       0       0       0
> rt$gP <- factor( floor((rt$P-1996)/3)*3+1996,
+                  labels=paste(1<-1996+0:6*3,'-',1+2,sep='') )
> with( rt, table( P, gP ) )
      gP
P      1996-1998 1999-2001 2002-2004 2005-2007 2008-2010 2011-2013 2014-2016
  1996     13126         0         0         0         0         0         0
  1997     13585         0         0         0         0         0         0
  1998     14127         0         0         0         0         0         0
  1999         0     14601         0         0         0         0         0
  2000         0     15097         0         0         0         0         0
  2001         0     15583         0         0         0         0         0
  2002         0         0     16131         0         0         0         0
  2003         0         0     16603         0         0         0         0
  2004         0         0     17061         0         0         0         0
  2005         0         0         0     17494         0         0         0
  2006         0         0         0     17924         0         0         0
  2007         0         0         0     18126         0         0         0
  2008         0         0         0         0     18191         0         0
  2009         0         0         0         0     18245         0         0
  2010         0         0         0         0     18283         0         0
  2011         0         0         0         0         0     18340         0
  2012         0         0         0         0         0     18353         0
  2013         0         0         0         0         0     18348         0
  2014         0         0         0         0         0         0     18360
  2015         0         0         0         0         0         0     18368
  2016         0         0         0         0         0         0     18404
> str( rt )
'data.frame':        354350 obs. of  18 variables:
 $ sex  : Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ state: Factor w/ 3 levels "noDM","T1","T2": 2 2 2 2 2 2 2 2 2 2 ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
 $ P    : num  1997 1999 1999 2000 2000 ...
 $ C    : num  1996 1998 1999 1999 1999 ...
 $ dur  : num  0.1 0.1 0.1 0.1 0.35 0.1 0.1 0.1 0.35 0.1 ...
 $ Dcvd : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dcan : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dres : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Doth : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y    : num  1.43e-04 2.26e-05 7.67e-05 3.79e-04 1.76e-04 ...
 $ T1   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Ax   : num  0.667 0.667 0.333 0.667 0.667 ...
 $ Px   : num  1997 1999 2000 2000 2000 ...
 $ Cx   : num  1997 1999 1999 2000 2000 ...
 $ gP   : Factor w/ 7 levels "1996-1998","1999-2001",..: 1 2 2 2 2 2 3 3 3 3 ...
```

A short overview of the number of events, deaths and person-years. This covers everyone in the population, including the persons with prevalent diabetes at 1996-01-01:

```
> mT <- addmargins( xtabs( cbind(T1,T2,DM=T1+T2,Dcvd,Dcan,Dres,Doth,D,Y) ~
+                          floor(P) + state + sex, data=rt ), 1:3 )
> mG <- addmargins( xtabs( cbind(T1,T2,DM=T1+T2,Dcvd,Dcan,Dres,Doth,D,Y) ~
+                               gP  + state + sex, data=rt ), 1:3 )
> str( mT ) ; str( mG )
 'table' num [1:22, 1:4, 1:3, 1:9] 678 684 657 592 596 586 602 545 509 517 ...
 - attr(*, "dimnames")=List of 4
  ..$ floor(P): chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ state   : chr [1:4] "noDM" "T1" "T2" "Sum"
  ..$ sex     : chr [1:3] "M" "W" "Sum"
  ..$         : chr [1:9] "T1" "T2" "DM" "Dcvd" ...

 'table' num [1:8, 1:4, 1:3, 1:9] 2019 1774 1656 1635 1643 ...
 - attr(*, "dimnames")=List of 4
  ..$ gP   : chr [1:8] "1996-1998" "1999-2001" "2002-2004" "2005-2007" ...
  ..$ state: chr [1:4] "noDM" "T1" "T2" "Sum"
  ..$ sex  : chr [1:3] "M" "W" "Sum"
  ..$      : chr [1:9] "T1" "T2" "DM" "Dcvd" ...

> length( mG ) ; dimnames( mG )

[1] 864

$gP
[1] "1996-1998" "1999-2001" "2002-2004" "2005-2007" "2008-2010" "2011-2013" "2014-2016"
[8] "Sum"

$state
[1] "noDM" "T1"   "T2"   "Sum"

$sex
[1] "M"   "W"   "Sum"

[[4]]
[1] "T1"   "T2"   "DM"   "Dcvd" "Dcan" "Dres" "Doth" "D"    "Y"

> zz <- fCtable( mG[,,"Sum",], row.vars=2:1, w=10, d=1 )
> zz[,1:8] <- gsub("\\.0","", zz[,1:8] )
> ftable(zz)
```

		T1	T2	DM	Dcvd	Dcan	Dres	Doth
state	gP							
noDM	1996-1998	3,478	33,986	37,464	58,088	42,494	15,170	43,916
	1999-2001	2,994	36,887	39,881	55,457	42,406	14,917	41,430
	2002-2004	2,816	49,185	52,001	50,897	40,506	15,295	41,733
	2005-2007	2,780	44,326	47,106	42,956	40,444	14,122	42,943
	2008-2010	2,734	56,453	59,187	36,812	39,304	15,031	44,762
	2011-2013	2,477	69,728	72,205	31,855	39,093	14,811	40,947
	2014-2016	2,433	53,387	55,820	29,588	38,229	14,137	41,226
	Sum	19,712	343,952	363,664	305,653	282,476	103,483	296,957
T1	1996-1998	.	.	.	868	290	137	671
	1999-2001	.	.	.	997	365	130	643
	2002-2004	.	.	.	835	333	175	860
	2005-2007	.	.	.	610	397	137	950
	2008-2010	.	.	.	501	402	183	803
	2011-2013	.	.	.	348	284	157	643
	2014-2016	.	.	.	243	237	113	450
	Sum	.	.	.	4,402	2,308	1,032	5,020
T2	1996-1998	.	.	.	8,133	2,788	1,114	3,564
	1999-2001	.	.	.	8,559	3,556	1,294	4,005
	2002-2004	.	.	.	8,084	4,025	1,860	5,199
	2005-2007	.	.	.	7,485	4,760	1,973	6,414
	2008-2010	.	.	.	7,080	5,508	2,562	7,416
	2011-2013	.	.	.	7,546	6,704	3,013	7,907
	2014-2016	.	.	.	7,850	7,965	3,584	9,052
	Sum	.	.	.	54,737	35,306	15,400	43,557
Sum	1996-1998	3,478	33,986	37,464	67,089	45,572	16,421	48,151
	1999-2001	2,994	36,887	39,881	65,013	46,327	16,341	46,078
	2002-2004	2,816	49,185	52,001	59,816	44,864	17,330	47,792
	2005-2007	2,780	44,326	47,106	51,051	45,601	16,232	50,307
	2008-2010	2,734	56,453	59,187	44,393	45,214	17,776	52,981
	2011-2013	2,477	69,728	72,205	39,749	46,081	17,981	49,497
	2014-2016	2,433	53,387	55,820	37,681	46,431	17,834	50,728
	Sum	19,712	343,952	363,664	364,792	320,090	119,915	345,534

```
> fCtable( mT[,-4,"Sum",], row.vars=2:1, w=7 )
```

		T1	T2	DM	Dcvd	Dcan	Dres	Doth	D	Y
state	floor(P)									
noDM	1996	1,194	11,405	12,599	19,669	14,187	5,178	15,229	54,263	5,194
	1997	1,173	10,757	11,930	19,557	14,220	5,079	14,592	53,448	5,209
	1998	1,111	11,824	12,935	18,862	14,087	4,913	14,095	51,957	5,221
	1999	1,005	12,446	13,451	19,035	14,193	5,239	13,927	52,394	5,230
	2000	988	12,197	13,185	18,024	14,121	4,786	13,711	50,642	5,241
	2001	1,001	12,244	13,245	18,398	14,092	4,892	13,792	51,174	5,255
	2002	988	15,356	16,344	17,839	13,560	5,315	14,221	50,935	5,265
	2003	931	16,819	17,750	17,153	13,433	5,227	13,931	49,744	5,270
	2004	897	17,010	17,907	15,905	13,513	4,753	13,581	47,752	5,273
	2005	896	14,684	15,580	15,093	13,523	4,620	13,593	46,829	5,279
	2006	936	14,112	15,048	14,252	13,684	4,570	14,110	46,616	5,289
	2007	948	15,530	16,478	13,611	13,237	4,932	15,240	47,020	5,304
	2008	913	17,400	18,313	12,590	13,240	4,879	14,957	45,666	5,324
	2009	925	18,482	19,407	12,416	12,968	5,231	14,970	45,585	5,342
	2010	896	20,571	21,467	11,806	13,096	4,921	14,835	44,658	5,354
	2011	854	28,743	29,597	10,929	13,244	4,935	13,752	42,860	5,362
	2012	801	22,799	23,600	10,721	13,105	4,822	13,451	42,099	5,366
	2013	822	18,186	19,008	10,205	12,744	5,054	13,744	41,747	5,378
	2014	806	17,241	18,047	9,879	12,655	4,551	13,391	40,476	5,403
	2015	851	17,625	18,476	10,050	12,808	4,841	13,815	41,514	5,438

	2016	776	18,521	19,297	9,659	12,766	4,745	14,020	41,190	5,460
	Sum	19,712	343,952	363,664	305,653	282,476	103,483	296,957	988,569	111,456
T1	1996	.	.	.	283	101	40	192	616	22
	1997	.	.	.	311	94	47	212	664	23
	1998	.	.	.	274	95	50	267	686	23
	1999	.	.	.	338	124	42	202	706	23
	2000	.	.	.	328	126	48	219	721	24
	2001	.	.	.	331	115	40	222	708	24
	2002	.	.	.	276	116	59	299	750	24
	2003	.	.	.	306	99	62	292	759	24
	2004	.	.	.	253	118	54	269	694	25
	2005	.	.	.	210	141	44	316	711	25
	2006	.	.	.	241	120	45	282	688	25
	2007	.	.	.	159	136	48	352	695	25
	2008	.	.	.	177	130	54	278	639	25
	2009	.	.	.	156	136	69	279	640	24
	2010	.	.	.	168	136	60	246	610	24
	2011	.	.	.	128	111	66	245	550	24
	2012	.	.	.	115	96	55	193	459	24
	2013	.	.	.	105	77	36	205	423	24
	2014	.	.	.	85	86	33	164	368	23
	2015	.	.	.	78	80	41	134	333	23
	2016	.	.	.	80	71	39	152	342	23
	Sum	.	.	.	4,402	2,308	1,032	5,020	12,762	501
T2	1996	.	.	.	2,662	849	391	1,174	5,076	65
	1997	.	.	.	2,717	934	369	1,199	5,219	71
	1998	.	.	.	2,754	1,005	354	1,191	5,304	77
	1999	.	.	.	2,863	1,133	418	1,300	5,714	83
	2000	.	.	.	2,818	1,197	403	1,375	5,793	90
	2001	.	.	.	2,878	1,226	473	1,330	5,907	96
	2002	.	.	.	2,715	1,236	537	1,710	6,198	104
	2003	.	.	.	2,761	1,339	692	1,738	6,530	113
	2004	.	.	.	2,608	1,450	631	1,751	6,440	124
	2005	.	.	.	2,540	1,551	599	2,038	6,728	133
	2006	.	.	.	2,578	1,599	640	2,021	6,838	140
	2007	.	.	.	2,367	1,610	734	2,355	7,066	148
	2008	.	.	.	2,260	1,673	698	2,377	7,008	157
	2009	.	.	.	2,386	1,876	955	2,471	7,688	167
	2010	.	.	.	2,434	1,959	909	2,568	7,870	179
	2011	.	.	.	2,475	2,060	937	2,510	7,982	195
	2012	.	.	.	2,518	2,307	1,014	2,618	8,457	213
	2013	.	.	.	2,553	2,337	1,062	2,779	8,731	224
	2014	.	.	.	2,537	2,654	1,146	2,827	9,164	232
	2015	.	.	.	2,693	2,628	1,204	3,021	9,546	239
	2016	.	.	.	2,620	2,683	1,234	3,204	9,741	247
	Sum	.	.	.	54,737	35,306	15,400	43,557	149,000	3,096

12.2.1 Modeling the transition rates

We model all the rates shown in the figure 12.1; we do this by means of a function. However, since we need different knot vectors we cannot refer to these by the same name; the prediction methods takes the object with the rates from the global environment.

Hence we construct the object `knuder`, a list of lists of lists of lists of vectors that holds the knots a for the three spline terms classified by `from`, `to` and `sex`. Note the assignment to `knuder` in the global environment using "`<<-`":

```
> fR <-
+ function( fr, to, nkn )
+ {
+ ml <- NULL
+ for( sx in levels(rt$sex) )
+    {
+ dat <- subset( rt, state==fr & sex==sx )
+ dat$D <- dat[,to]
+ a.kn <- c(20,with( dat, quantile( rep(Ax,D), (1:nkn[1]-0.5)/nkn[1] ) ))
+ p.kn <-      with( dat, quantile( rep(Px,D), (1:nkn[2]-0.5)/nkn[2] ) )
+ c.kn <-      with( dat, quantile( rep(Cx,D), (1:nkn[3]-0.5)/nkn[3] ) )
+ knuder[[fr]][[to]][[sx]] <<- list( A=a.kn, P=p.kn, C=c.kn )
+ form <- paste0("D ~ Ns(Ax,knots=knuder$",fr,"$",to,"$",sx,"$A) + ",
+                    "Ns(Px,knots=knuder$",fr,"$",to,"$",sx,"$P) + ",
+                  "Ns(Px-Ax,knots=knuder$",fr,"$",to,"$",sx,"$C)" )
+ ml <- c( ml, list( glm( formula = form,
+                         offset = log(Y),
+                         family = poisson,
+                         data = dat ) ) )
+    }
+ names(ml) <- levels(rt$sex)
+ ml
+ }
```

With this function we can now fit the three all cause mortality rates and inspect the resulting `knuder` object:

```
> knuder <- list()
> nDtot <- fR("noDM","D",c(6,4,6))
> T1tot <- fR(  "T1","D",c(6,4,6))
> T2tot <- fR(  "T2","D",c(6,4,6))
> str( knuder )
List of 3
 $ noDM:List of 1
  ..$ D:List of 2
  .. ..$ M:List of 3
  .. .. ..$ A: Named num [1:7] 20 50.7 65.3 73.3 78.7 ...
  .. .. .. ..- attr(*, "names")= chr [1:7] "" "8.333333%" "25%" "41.66667%" ...
  .. .. ..$ P: Named num [1:4] 1998 2003 2008 2014
  .. .. .. ..- attr(*, "names")= chr [1:4] "12.5%" "37.5%" "62.5%" "87.5%"
  .. .. ..$ C: Named num [1:6] 1914 1921 1927 1933 1942 ...
  .. .. .. ..- attr(*, "names")= chr [1:6] "8.333333%" "25%" "41.66667%" "58.33333%" ...
  .. ..$ W:List of 3
  .. .. ..$ A: Named num [1:7] 20 57.7 71.7 79.3 84.3 ...
  .. .. .. ..- attr(*, "names")= chr [1:7] "" "8.333333%" "25%" "41.66667%" ...
  .. .. ..$ P: Named num [1:4] 1998 2003 2008 2014
  .. .. .. ..- attr(*, "names")= chr [1:4] "12.5%" "37.5%" "62.5%" "87.5%"
  .. .. ..$ C: Named num [1:6] 1910 1917 1922 1927 1935 ...
  .. .. .. ..- attr(*, "names")= chr [1:6] "8.333333%" "25%" "41.66667%" "58.33333%" ...
 $ T1  :List of 1
  ..$ D:List of 2
  .. ..$ M:List of 3
  .. .. ..$ A: Named num [1:7] 20 44.3 56.7 64.3 70.7 ...
  .. .. .. ..- attr(*, "names")= chr [1:7] "" "8.333333%" "25%" "41.66667%" ...
  .. .. ..$ P: Named num [1:4] 1999 2003 2008 2013
  .. .. .. ..- attr(*, "names")= chr [1:4] "12.5%" "37.5%" "62.5%" "87.5%"
  .. .. ..$ C: Named num [1:6] 1919 1927 1935 1942 1949 ...
  .. .. .. ..- attr(*, "names")= chr [1:6] "8.333333%" "25%" "41.66667%" "58.33333%" ...
```

```
  .. ..$ W:List of 3
  .. .. ..$ A: Named num [1:7] 20 49.7 64.3 72.7 78.3 ...
  .. .. .. ..- attr(*, "names")= chr [1:7] "" "8.333333%" "25%" "41.66667%" ...
  .. .. ..$ P: Named num [1:4] 1998 2003 2007 2012
  .. .. .. ..- attr(*, "names")= chr [1:4] "12.5%" "37.5%" "62.5%" "87.5%"
  .. .. ..$ C: Named num [1:6] 1914 1921 1926 1933 1942 ...
  .. .. .. ..- attr(*, "names")= chr [1:6] "8.333333%" "25%" "41.66667%" "58.33333%" ...
 $ T2  :List of 1
  ..$ D:List of 2
  .. ..$ M:List of 3
  .. .. ..$ A: Named num [1:7] 20 59.3 68.3 73.7 78.7 ...
  .. .. .. ..- attr(*, "names")= chr [1:7] "" "8.333333%" "25%" "41.66667%" ...
  .. .. ..$ P: Named num [1:4] 2000 2006 2011 2015
  .. .. .. ..- attr(*, "names")= chr [1:4] "12.5%" "37.5%" "62.5%" "87.5%"
  .. .. ..$ C: Named num [1:6] 1916 1924 1929 1935 1941 ...
  .. .. .. ..- attr(*, "names")= chr [1:6] "8.333333%" "25%" "41.66667%" "58.33333%" ...
  .. ..$ W:List of 3
  .. .. ..$ A: Named num [1:7] 20 63.3 73.7 79.3 83.7 ...
  .. .. .. ..- attr(*, "names")= chr [1:7] "" "8.333333%" "25%" "41.66667%" ...
  .. .. ..$ P: Named num [1:4] 1999 2005 2010 2015
  .. .. .. ..- attr(*, "names")= chr [1:4] "12.5%" "37.5%" "62.5%" "87.5%"
  .. .. ..$ C: Named num [1:6] 1912 1919 1924 1929 1935 ...
  .. .. .. ..- attr(*, "names")= chr [1:6] "8.333333%" "25%" "41.66667%" "58.33333%" ...
```

With this function we can fit the further models for the 24 cause-specific mortality rates:

```
> nDT1  <- fR("noDM",  "T1",c(6,4,6))
> nDT2  <- fR("noDM",  "T2",c(6,4,6))
> nDcvd <- fR("noDM","Dcvd",c(6,4,6))
> nDcan <- fR("noDM","Dcan",c(6,4,6))
> nDres <- fR("noDM","Dres",c(6,4,6))
> nDoth <- fR("noDM","Doth",c(6,4,6))
> T1cvd <- fR(  "T1","Dcvd",c(6,4,6))
> T1can <- fR(  "T1","Dcan",c(6,4,6))
> T1res <- fR(  "T1","Dres",c(6,4,6))
> T1oth <- fR(  "T1","Doth",c(6,4,6))
> T2cvd <- fR(  "T2","Dcvd",c(6,4,6))
> T2can <- fR(  "T2","Dcan",c(6,4,6))
> T2res <- fR(  "T2","Dres",c(6,4,6))
> T2oth <- fR(  "T2","Doth",c(6,4,6))
```

We have now models for all 14 transitions for each sex separately, plus — as a check — models for the overall mortality which should yield predictions very close to the sum of the predictions from the separate causes of death. They should not necessarily be identical because we have used different parametrizations.

We have the knots in the object `knuder`, where each of the listed lists is a list of lists with names `M` and `W` which again are lists with names `A`, `P` and `C`:

```
> lapply( knuder, names )
```

As a prerequisite for subsequent plotting we create 8 shades of blue and red, respectively (the first color generated is sooo weak):

```
> library( RColorBrewer )
> clb <- brewer.pal(9,"Blues")[-1]
> clr <- brewer.pal(9,"Reds" )[-1]
> # plot( rep(1:8,2), rep(1:2,each=8), col=c(clb,clr),
> #       pch=15, cex=15, ylim=c(0,3) )
```

Finally we save the fitted models and the object with the knots, `knuder`:

```
> save( knuder,
+ nDtot, T1tot, T2tot,
+ nDT1 , nDT2 ,
+ nDcvd, nDcan, nDres, nDoth,
+ T1cvd, T1can, T1res, T1oth,
+ T2cvd, T2can, T2res, T2oth,
+ file="../nydata/trmods.Rda" )


> load( file="../nydata/trmods.Rda" )
```

Rates by age at different dates

We can plot the age-specific rates by cause for T1, T2 and noDM separately:

```
> nd <- data.frame( Ax=0:90, Y=1 )
> getr <- function( mod ){
+ rates <- NULL
+ for( p in seq(1996,2017,3) )
+      rates <- cbind( rates,
+                      ci.pred( mod, cbind(nd,Px=p) ) )
+ rates
+ }
> ( whclr <- rep(1:8,each=2) + rep(0:1,8)*8 )
 [1]  1  9  2 10  3 11  4 12  5 13  6 14  7 15  8 16
> ( whclm <- as.vector( outer(1:3,whclr-1,function(x,y) x+y*3) ) )
 [1]  1  2  3 25 26 27  4  5  6 28 29 30  7  8  9 31 32 33 10 11 12 34 35 36 13 14 15 37
[29] 38 39 16 17 18 40 41 42 19 20 21 43 44 45 22 23 24 46 47 48
> plmod <- function( mod ) {
+ plot( NA, log="y", ylim=c(0.005,100), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
+ abline( h=outer(c(1,2,5),-3:2,function(x,y)x*10^y),
+         v=seq(10,90,10), col=gray(0.8) )
+ matshade( nd$Ax, cbind( getr( mod$M ),
+                         getr( mod$W ) )[,whclm],
+           col=c(clb,clr)[whclr], lwd=1 )
+ }
> a1 <- function(){
+ axis( side=1, at=seq(0,80,20) )
+ axis( side=1, at=seq(0,90,10), labels=NA, tcl=-0.4 )
+ axis( side=1, at=seq(0,95, 5), labels=NA, tcl=-0.2 )
+ }
> a2 <- function(){
+ lpos <- outer(c(1,2,5),-3:2,function(x,y)x*10^y)
+ axis( side=2, at=lpos, labels=formatC(lpos,format="g",dig=3) )
+ axis( side=2, at=outer(1:9,-3:2,function(x,y)x*10^y), labels=NA, tcl=-0.3 )
+ }
> par( mar=c(0,0,0,0), oma=c(4,4,2,1), mfcol=c(5,3),
+     las=1, bty="o" )
> plmod( T1cvd ) ; a2()
> plmod( T1can ) ; a2()
> plmod( T1res ) ; a2()
> plmod( T1oth ) ; a2()
```

```
> plmod( T1tot ) ; a2() ; a1()
> plmod( T2cvd )
> plmod( T2can )
> plmod( T2res )
> plmod( T2oth )
> plmod( T2tot ) ; a1()
> plmod( nDcvd )
> plmod( nDcan )
> plmod( nDres )
> plmod( nDoth )
> plmod( nDtot ) ; a1()
> mtext( rep("Age",3)          , side=1, at=(1:3*2-1)/6, line=2, cex=0.66, outer=T )
> mtext( c("All causes","Other","Respir","Cancer","CVD"),
+                               side=2, at=(1:5*2-1)/10, line=3, cex=0.66, outer=T, las=0 )
> mtext( c("T1D","T2D","no DM"), side=3, at=(1:3*2-1)/6, line=1, cex=0.66, outer=T )
```

From figure 12.2 it is seen that it is particularly the mortality rates for ”Other” that differs between DM and non-DM.

Rates by date for different ages

We can plot the rates in different ages versus date by cause for T1, T2 and noDM separately:

```
> np <- data.frame( Px=seq(1996,2017,,100), Y=1 )
> getp <- function( mod ){
+     rates <- NULL
+ for( a in 2:9*10 )
+      rates <- cbind( rates,
+                      ci.pred( mod, cbind(np,Ax=a) ) )
+ rates
+ }
> ( whclr <- rep(1:8,each=2) + rep(0:1,8)*8 )
 [1]  1  9  2 10  3 11  4 12  5 13  6 14  7 15  8 16
> ( whclm <- as.vector( outer(1:3,whclr-1,function(x,y) x+y*3) ) )
 [1]  1  2  3 25 26 27  4  5  6 28 29 30  7  8  9 31 32 33 10 11 12 34 35 36 13 14 15 37
[29] 38 39 16 17 18 40 41 42 19 20 21 43 44 45 22 23 24 46 47 48
> plmod <- function( mod ) {
+ plot( NA, log="y", ylim=c(0.005,250), yaxt="n",
+                    xlim=c(1996,2017), xaxt="n" )
+ abline( h=outer(c(1,2,5),-3:2,function(x,y)x*10^y),
+         v=seq(10,90,10), col=gray(0.8) )
+ matshade( np$Px, cbind( getp( mod$M ),
+                         getp( mod$W ) )[,whclm],
+           col=c(clb,clr)[whclr], lwd=1 )
+ }
> a1p <- function(){
+ axis( side=1, at=seq(1995,2020,5) )
+ axis( side=1, at=seq(1995,2020,2), labels=NA, tcl=-0.4 )
+ axis( side=1, at=seq(1995,2020,1), labels=NA, tcl=-0.2 )
+ }
> par( mar=c(0,0,0,0), oma=c(4,4,2,1), mfcol=c(5,3),
+     las=1, bty="o" )
> plmod( T1cvd ) ; a2()
> plmod( T1can ) ; a2()
> plmod( T1res ) ; a2()
> plmod( T1oth ) ; a2()
```

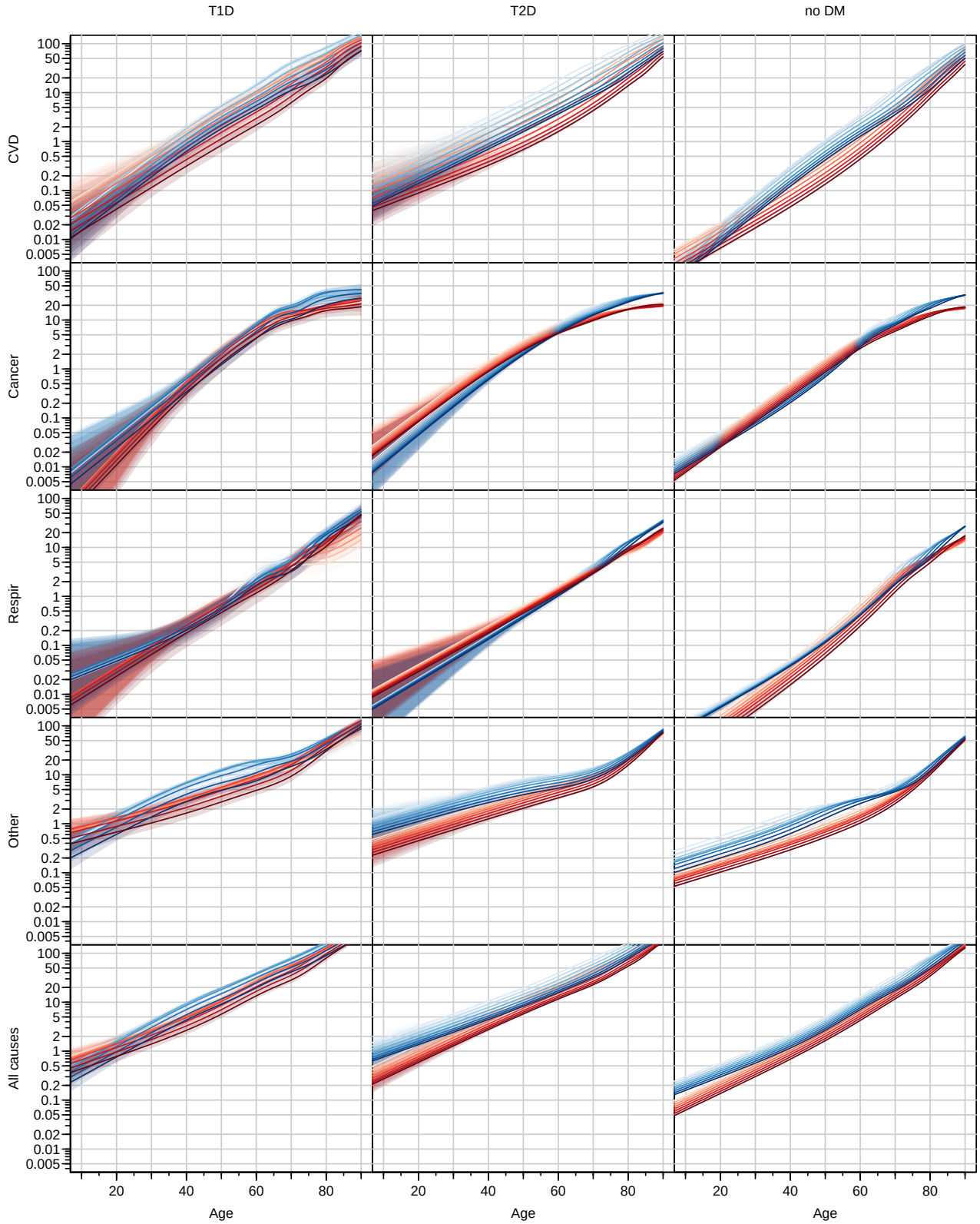


Figure 12.2: *Cause-specific mortality rates for T1, T2 and noDM. Cross sectional rates for the dates 1996, 1999,. . . ,2017, colouring from light to dark. Blue is for men, red for women.* `./graph/yll-cod-a-rates`

```
> plmod( T1tot ) ; a2() ; a1p()
> plmod( T2cvd )
> plmod( T2can )
> plmod( T2res )
> plmod( T2oth )
> plmod( T2tot ) ; a1p()
> plmod( nDcvd )
> plmod( nDcan )
> plmod( nDres )
> plmod( nDoth )
> plmod( nDtot ) ; a1p()
> mtext( rep("Date of FU",3)   , side=1, at=(1:3*2-1)/6, line=2, cex=0.66, outer=T )
> mtext( c("All causes","Other","Respir","Cancer","CVD"),
+                                side=2, at=(1:5*2-1)/10, line=3, cex=0.66, outer=T, las=0 )
> mtext( c("T1D","T2D","no DM"), side=3, at=(1:3*2-1)/6, line=1, cex=0.66, outer=T )
```

From figure 12.2 it is seen that it is particularly the mortality rates for ”Other” that differs between DM and non-DM.

12.2.2 Checking with all-cause mortality

We also fitted models for all-cause mortality, and these should hopefully be similar to the sum of the predictions from the separate causes of death. We do the comparison for the dates 1999, 2005, 2011 and 2017 only to make plots simpler.

```
> par( mar=c(0,0,0,0), oma=c(4,4,2,1), mfcol=c(3,2),
+     las=1, bty="o" )
> ( wh  <- 1+0:3*6 )
[1]  1  7 13 19
> ( wht <- as.vector( outer( 0:2, wh, "+" ) ) )
 [1]  1  2  3  7  8  9 13 14 15 19 20 21
> # T1 men
> plot( NA, log="y", ylim=c(0.05,500), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=outer(c(1,2,5),-3:2,function(x,y)x*10^y),
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax, getr( T1cvd$M )[,wh]+
+                  getr( T1can$M )[,wh]+
+                  getr( T1res$M )[,wh]+
+                  getr( T1oth$M )[,wh],
+           col=clb[c(2,4,6,8)],
+           lwd=3, lty="31", lend="butt" )
> matshade( nd$Ax, getr( T1tot$M )[,wht],
+           col=clb[c(2,4,6,8)], lwd=2, lty=1 )
> a2()
> # T2 men
> plot( NA, log="y", ylim=c(0.05,500), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=outer(c(1,2,5),-3:2,function(x,y)x*10^y),
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax, getr( T2cvd$M )[,wh]+
+                  getr( T2can$M )[,wh]+
+                  getr( T2res$M )[,wh]+
+                  getr( T2oth$M )[,wh],
+           col=clb[c(2,4,6,8)], lwd=3, lty="31", lend="butt" )
```

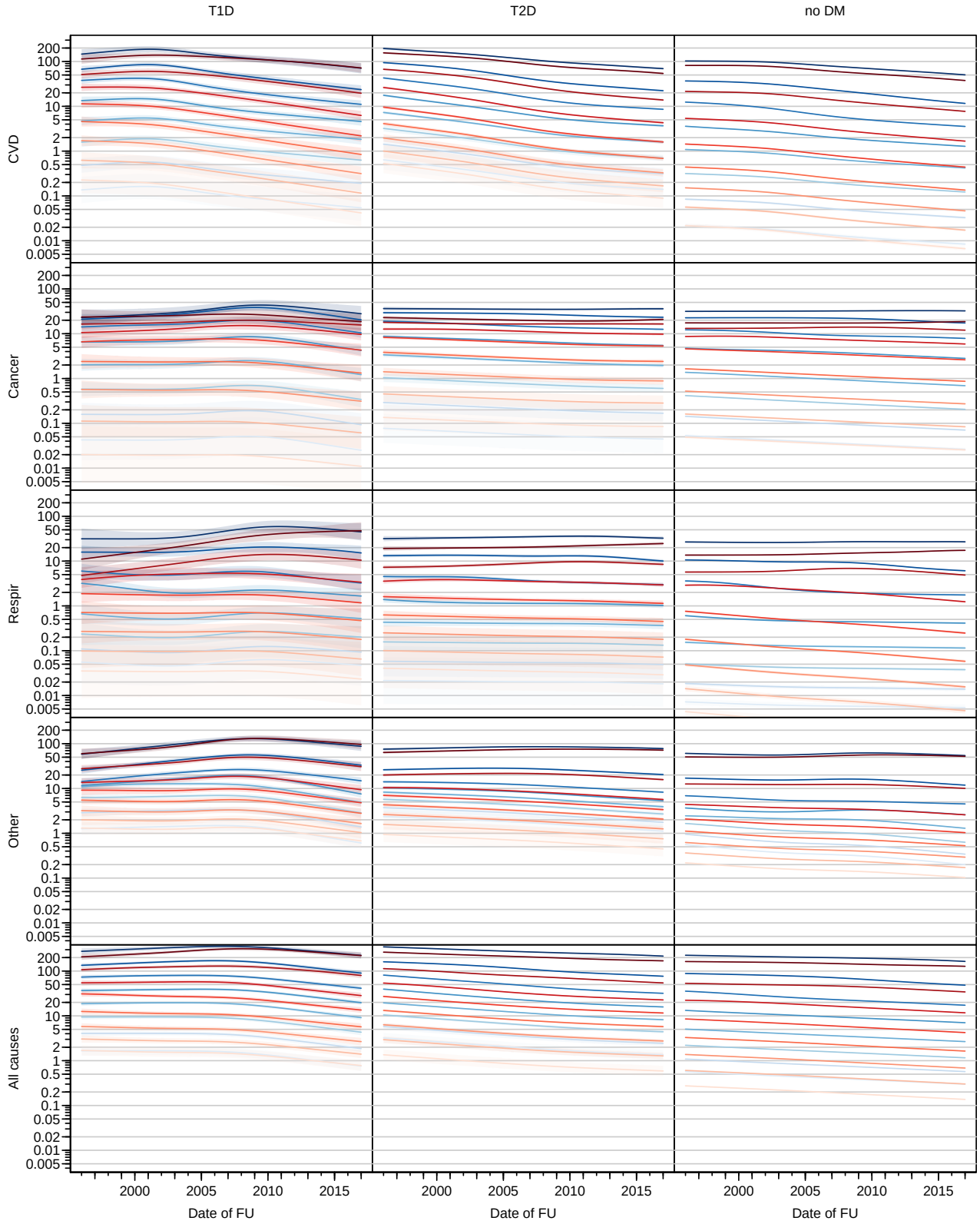


Figure 12.3: *Cause-specific mortality rates for T1, T2 and noDM. Age-specific rates in ages 20, 30, . . . , 90; colouring from light (youngest) to dark (oldest). Blue is for men, red for women.* `./graph/yll-cod-p-rates`

```
> matshade( nd$Ax, getr( T2tot$M )[,wht],
+           col=clb[c(2,4,6,8)], lwd=2, lty=1 )
> a2()
> # nD men
> plot( NA, log="y", ylim=c(0.05,500), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=outer(c(1,2,5),-3:2,function(x,y)x*10^y),
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax, getr( nDcvd$M )[,wh]+
+                  getr( nDcan$M )[,wh]+
+                  getr( nDres$M )[,wh]+
+                  getr( nDoth$M )[,wh],
+           col=clb[c(2,4,6,8)], lwd=3, lty="31", lend="butt" )
> matshade( nd$Ax, getr( nDtot$M )[,wht],
+           col=clb[c(2,4,6,8)], lwd=2, lty=1 )
> a2() ; a1()
> # T1 women
> plot( NA, log="y", ylim=c(0.05,500), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=outer(c(1,2,5),-3:2,function(x,y)x*10^y),
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax, getr( T1cvd$W )[,wh]+
+                  getr( T1can$W )[,wh]+
+                  getr( T1res$W )[,wh]+
+                  getr( T1oth$W )[,wh],
+           col=clr[c(2,4,6,8)], lwd=3, lty="31", lend="butt" )
> matshade( nd$Ax, getr( T1tot$W )[,wht],
+           col=clr[c(2,4,6,8)], lwd=2, lty=1 )
> # T2 women
> plot( NA, log="y", ylim=c(0.05,500), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=outer(c(1,2,5),-3:2,function(x,y)x*10^y),
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax, getr( T2cvd$W )[,wh]+
+                  getr( T2can$W )[,wh]+
+                  getr( T2res$W )[,wh]+
+                  getr( T2oth$W )[,wh],
+           col=clr[c(2,4,6,8)], lwd=3, lty="31", lend="butt" )
> matshade( nd$Ax, getr( T2tot$W )[,wht],
+           col=clr[c(2,4,6,8)], lwd=2, lty=1 )
> # nD women
> plot( NA, log="y", ylim=c(0.05,500), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=outer(c(1,2,5),-3:2,function(x,y)x*10^y),
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax, getr( nDcvd$W )[,wh]+
+                  getr( nDcan$W )[,wh]+
+                  getr( nDres$W )[,wh]+
+                  getr( nDoth$W )[,wh],
+           col=clr[c(2,4,6,8)], lwd=3, lty="31", lend="butt" )
> matshade( nd$Ax, getr( nDtot$W )[,wht],
+           col=clr[c(2,4,6,8)], lwd=2, lty=1 )
> a1()
> mtext( rep("Age",2)          , side=1, at=c(1,3  )/4, line=2, cex=0.66, outer=T )
> mtext( c("T1D","T2D","no DM"), side=2, at=c(5,3,1)/6, line=2, cex=0.66, outer=T, las=0 )
```

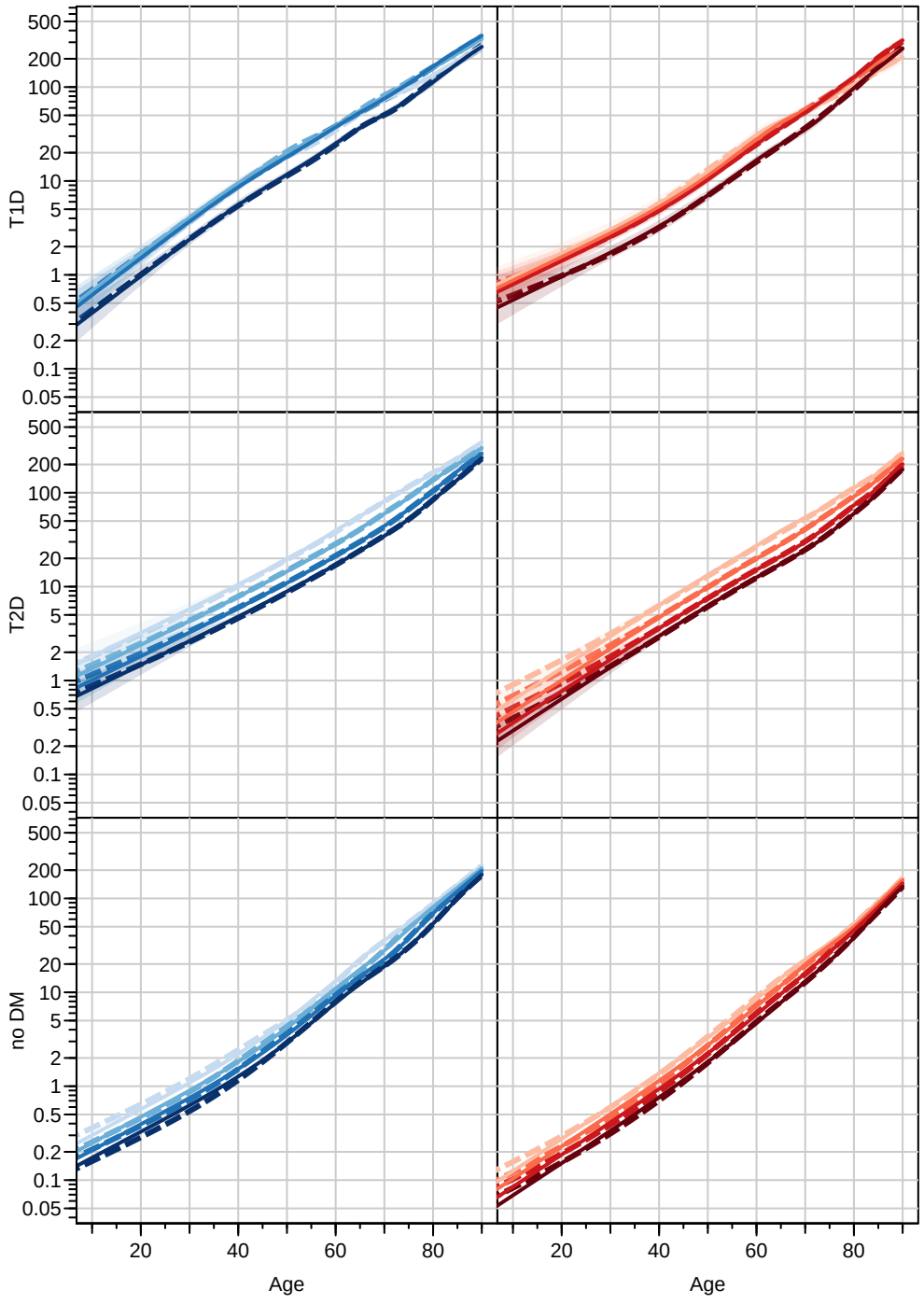


Figure 12.4: *All cause mortality rates for T1, T2 and noDM (top down). Cross sectional rates for the dates 1999, 2005, 2011,2017, colouring from light to dark. Broken lines are sum of the predicted rates for the four separate analyses of rates for the four causes of death, full lines (and shades) are from the models for all cause mortality. Blue is for men, red for women.* `./graph/yll-cod-tot-rates`

In a similar plot we the ratio of the sum of the four cause-specific rates to the estimated total mortality rates and

```
> par( mar=c(0,0,0,0), oma=c(4,4,2,1), mfcol=c(3,2),
+     las=1, bty="o" )
> ( wh  <- 1+0:3*6 )
[1]  1  7 13 19
> ( wht <- as.vector( outer( 0:2, wh, "+" ) ) )
 [1]  1  2  3  7  8  9 13 14 15 19 20 21
> # T1 men
> plot( NA, log="y", ylim=c(0.5,2), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=5:20/10,
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax,(getr( T1cvd$M )[,wh]+
+                  getr( T1can$M )[,wh]+
+                  getr( T1res$M )[,wh]+
+                  getr( T1oth$M )[,wh])/getr( T1tot$M )[,wh],
+           col=clb[c(2,4,6,8)],
+           lwd=3, lty=1, lend="butt" )
> a2()
> # T2 men
> plot( NA, log="y", ylim=c(0.5,2), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=5:20/10,
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax,(getr( T2cvd$M )[,wh]+
+                  getr( T2can$M )[,wh]+
+                  getr( T2res$M )[,wh]+
+                  getr( T2oth$M )[,wh])/getr( T2tot$M )[,wh],
+           col=clb[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
> a2()
> # nD men
> plot( NA, log="y", ylim=c(0.5,2), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=5:20/10,
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax,(getr( nDcvd$M )[,wh]+
+                  getr( nDcan$M )[,wh]+
+                  getr( nDres$M )[,wh]+
+                  getr( nDoth$M )[,wh])/getr( nDtot$M )[,wh],
+           col=clb[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
> a2() ; a1()
> # T1 women
> plot( NA, log="y", ylim=c(0.5,2), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=5:20/10,
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax,(getr( T1cvd$W )[,wh]+
+                  getr( T1can$W )[,wh]+
+                  getr( T1res$W )[,wh]+
+                  getr( T1oth$W )[,wh])/getr( T1tot$W )[,wh],
+           col=clr[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
> # T2 women
> plot( NA, log="y", ylim=c(0.5,2), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
```

```
> abline( h=5:20/10,
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax,(getr( T2cvd$W )[,wh]+
+                  getr( T2can$W )[,wh]+
+                  getr( T2res$W )[,wh]+
+                  getr( T2oth$W )[,wh])/getr( T2tot$W )[,wh],
+           col=clr[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
> # nD women
> plot( NA, log="y", ylim=c(0.5,2), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=5:20/10,
+         v=seq(10,90,10), col=gray(0.8) )
> matlines( nd$Ax,(getr( nDcvd$W )[,wh]+
+                  getr( nDcan$W )[,wh]+
+                  getr( nDres$W )[,wh]+
+                  getr( nDoth$W )[,wh])/getr( nDtot$W )[,wh],
+           col=clr[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
> a1()
> mtext( rep("Age",2)           , side=1, at=  c(1,3)/4, line=2, cex=0.66, outer=T )
> mtext( c("T1D","T2D","no DM"), side=2, at=  c(5,3,1)/6, line=1, cex=0.66, outer=T, las=0 )
```

Finally we plot the differences between the sum of the cause-specific rates and the estimated all cause mortality rates:

```
> par( mar=c(0,0,0,0), oma=c(4,4,2,1), mfcol=c(3,2),
+     las=1, bty="o" )
> ( wh  <- 1+0:3*6 )
[1]  1  7 13 19
> ( wht <- as.vector( outer( 0:2, wh, "+" ) ) )
 [1]  1  2  3  7  8  9 13 14 15 19 20 21
> A2 <- function(){
+ axis( side=2, at=-5:2*10 )
+ axis( side=2, at=-10:4*5, tcl= -0.3)
+ }
> # T1 men
> plot( NA,  ylim=c(-10,10), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=-10:10, v=seq(10,90,10), col=gray(0.8) )
> abline( h=0 )
> matlines( nd$Ax,(getr( T1cvd$M )[,wh]+
+                  getr( T1can$M )[,wh]+
+                  getr( T1res$M )[,wh]+
+                  getr( T1oth$M )[,wh])-getr( T1tot$M )[,wh],
+           col=clb[c(2,4,6,8)],
+           lwd=3, lty=1, lend="butt" )
> A2()
> # T2 men
> plot( NA,  ylim=c(-10,10), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=-10:10, v=seq(10,90,10), col=gray(0.8) )
> abline( h=0 )
> matlines( nd$Ax,(getr( T2cvd$M )[,wh]+
+                  getr( T2can$M )[,wh]+
+                  getr( T2res$M )[,wh]+
+                  getr( T2oth$M )[,wh])-getr( T2tot$M )[,wh],
+           col=clb[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
```

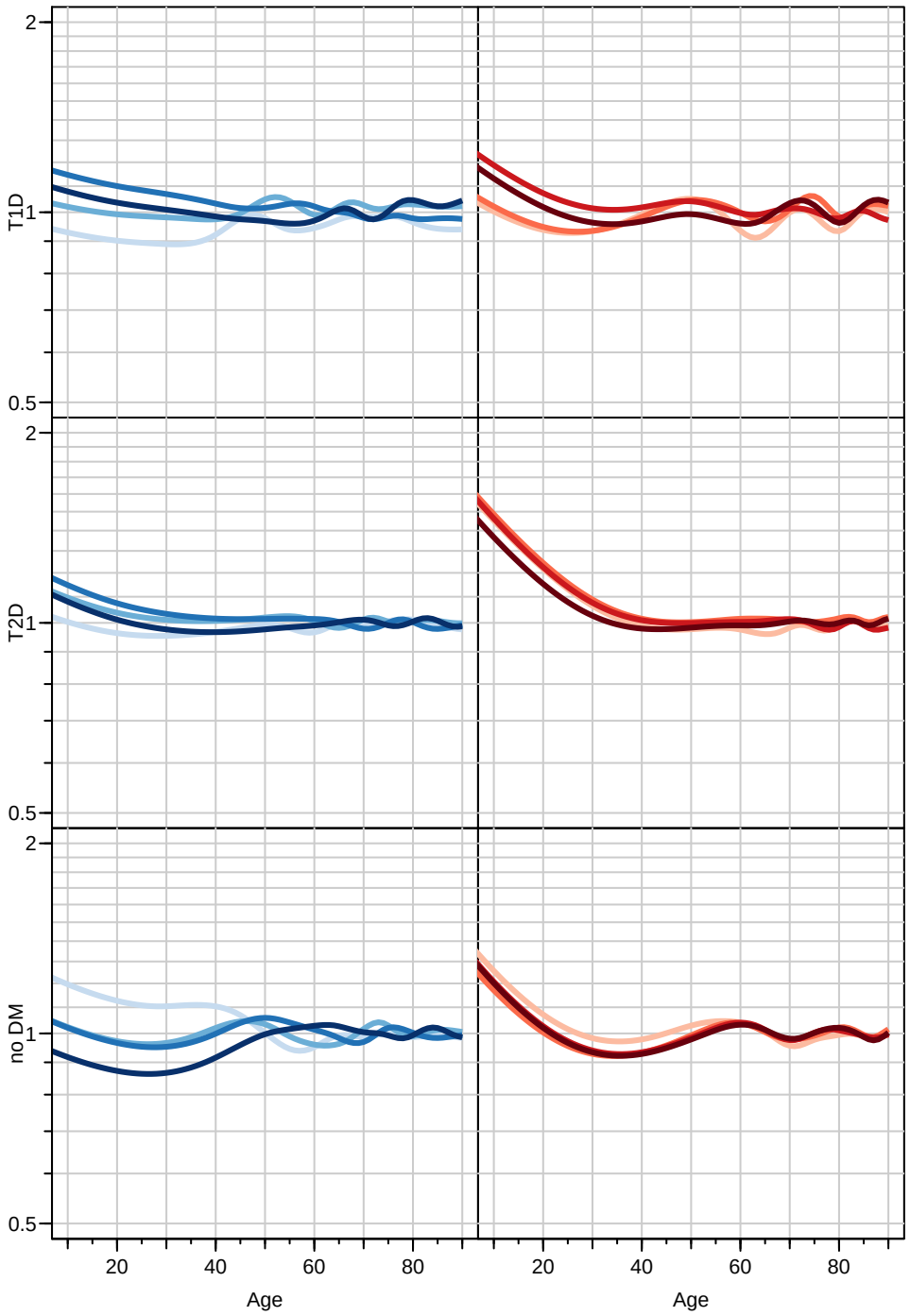


Figure 12.5: *Ratio of the sum of the cause-specific mortalities to the all cause mortality for the dates 1999, 2005, 2011,2017, colouring from light to dark. Blue is for men, red for women.* `./graph/yll-cod-tot-Rrates`

```
> A2()
> # nD men
> plot( NA,  ylim=c(-10,10), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=-10:10, v=seq(10,90,10), col=gray(0.8) )
> abline( h=0 )
> matlines( nd$Ax,(getr( nDcvd$M )[,wh]+
+                  getr( nDcan$M )[,wh]+
+                  getr( nDres$M )[,wh]+
+                  getr( nDoth$M )[,wh])-getr( nDtot$M )[,wh],
+           col=clb[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
> A2() ; a1()
> # T1 women
> plot( NA,  ylim=c(-10,10), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=-10:10, v=seq(10,90,10), col=gray(0.8) )
> abline( h=0 )
> matlines( nd$Ax,(getr( T1cvd$W )[,wh]+
+                  getr( T1can$W )[,wh]+
+                  getr( T1res$W )[,wh]+
+                  getr( T1oth$W )[,wh])-getr( T1tot$W )[,wh],
+           col=clr[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
> # T2 women
> plot( NA,  ylim=c(-10,10), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=-10:10, v=seq(10,90,10), col=gray(0.8) )
> abline( h=0 )
> matlines( nd$Ax,(getr( T2cvd$W )[,wh]+
+                  getr( T2can$W )[,wh]+
+                  getr( T2res$W )[,wh]+
+                  getr( T2oth$W )[,wh])-getr( T2tot$W )[,wh],
+           col=clr[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
> # nD women
> plot( NA,  ylim=c(-10,10), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
> abline( h=-10:10, v=seq(10,90,10), col=gray(0.8) )
> abline( h=0 )
> matlines( nd$Ax,(getr( nDcvd$W )[,wh]+
+                  getr( nDcan$W )[,wh]+
+                  getr( nDres$W )[,wh]+
+                  getr( nDoth$W )[,wh])-getr( nDtot$W )[,wh],
+           col=clr[c(2,4,6,8)], lwd=3, lty=1, lend="butt" )
> a1()
> mtext( rep("Age",2)          , side=1, at=  c(1,3)/4, line=2, cex=0.66, outer=T )
> mtext( c("T1D","T2D","no DM"), side=2, at=  c(5,3,1)/6, line=2, cex=0.66, outer=T, las=0 )
```

From figures 12.4, 12.5 and 12.6 we see a fine concordance between the sum of the fitted rates for the four causes and the fitted rates for the four causes together.

12.2.3 Comparing mortality rates

We then compare the cause-specific mortality rates between the three groups:

```
> plcmp <- function( m1, m0 ) {
+ plot( NA, log="y", ylim=c(0.1,100), yaxt="n",
+                    xlim=c(10,90), xaxt="n" )
+ abline( h=outer(c(1,2,5),-2:2,function(x,y)x*10^y),
```

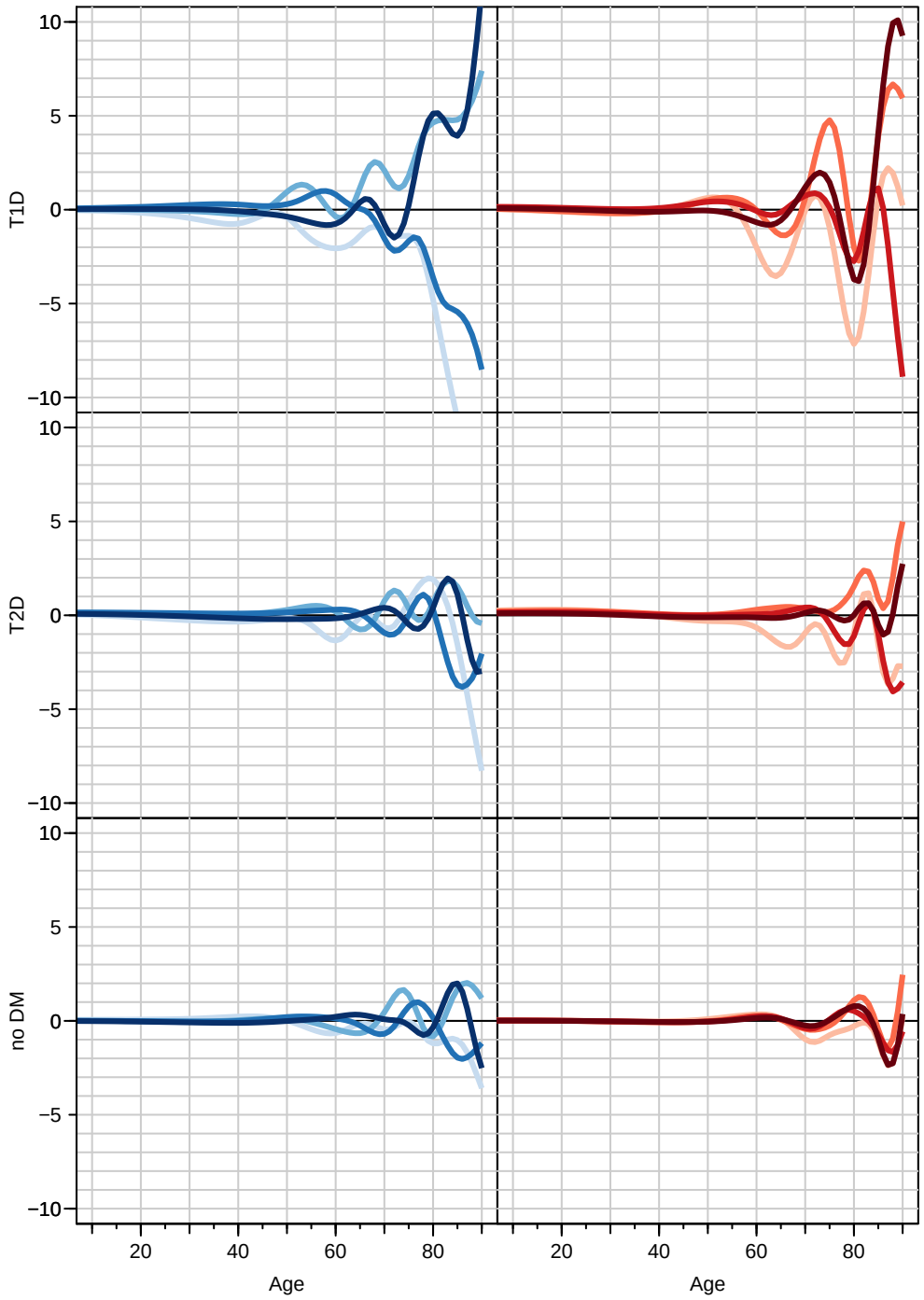


Figure 12.6: *Differences between the sum of the cause-specific mortalities and the all cause mortality (per 1000 PY) for the dates 1999, 2005, 2011, 2017, colouring from light to dark. Blue is for men, red for women.* `./graph/yll-cod-tot-Drates`

```
+         v=seq(10,90,10), col=gray(0.8) )
+ r1 <- cbind( getr( m1$M ), getr( m1$W ) )
+ r0 <- cbind( getr( m0$M ), getr( m0$W ) )
+ nr <- ncol(r1)/3
+ rr <- NULL
+ for( i in 1:nr ) rr <- cbind( rr, ci.ratio( r1[,(i-1)*3+1:3],
+                                              r0[,(i-1)*3+1:3] ) )
+ matshade( nd$Ax, rr[,whclm],
+           col=c(clb,clr)[whclr], lwd=2 )
+ abline( h=1 )
+ }
> par( mar=c(0,0,0,0), oma=c(4,4,2,1), mfcol=c(4,3),
+     las=1, bty="o" )
> plcmp( T1cvd, nDcvd ) ; a2()
> plcmp( T1can, nDcan ) ; a2()
> plcmp( T1res, nDres ) ; a2()
> plcmp( T1oth, nDoth ) ; a2() ; a1()
> plcmp( T2cvd, nDcvd )
> plcmp( T2can, nDcan )
> plcmp( T2res, nDres )
> plcmp( T2oth, nDoth ) ; a1()
> plcmp( T1cvd, T2cvd )
> plcmp( T1can, T2can )
> plcmp( T1res, T2res )
> plcmp( T1oth, T2oth ) ; a1()
> mtext( rep("Age",3), side=1, at=c(1,3,5)/6  , line=2, cex=0.66, outer=T )
> mtext( c("Other","Respir","Cancer","CVD"),
+        side=2, at=c(1,3,5,7)/8, line=2, cex=0.66, outer=T, las=0 )
> mtext( c("T1D vs no DM",
+          "T2D vs no DM",
+          "T1D vs T2D"),
+        side=3, at=c(1,3,5)/6  , line=1, cex=0.66, outer=T )
```

12.3 Transition probabilities

12.3.1 Set-up for YLL

We now set up state-transition matrices for each sex and date (the date where we take cross-sectional rates to compute state occupancy probabilities). The last two dimensions of the array from the transition matrix, at a given time — the third last dimension:

```
> int <- 1/12
> a.pt <- seq(int,100,int) - int/2
> states <- c("noDM","T1D","T2D","D-CVD","D-Can","D-Res","D-Oth")
> Tr <- ZArray( list( sex = levels(rt$sex),
+                    date = 1996:2017,
+                     age = a.pt,
+                    from = states,
+                      to = states ) )
> fC( length(Tr) ) ; str( Tr )
[1] "2,587,200"
 num [1:2, 1:22, 1:1200, 1:7, 1:7] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex : chr [1:2] "M" "W"
```

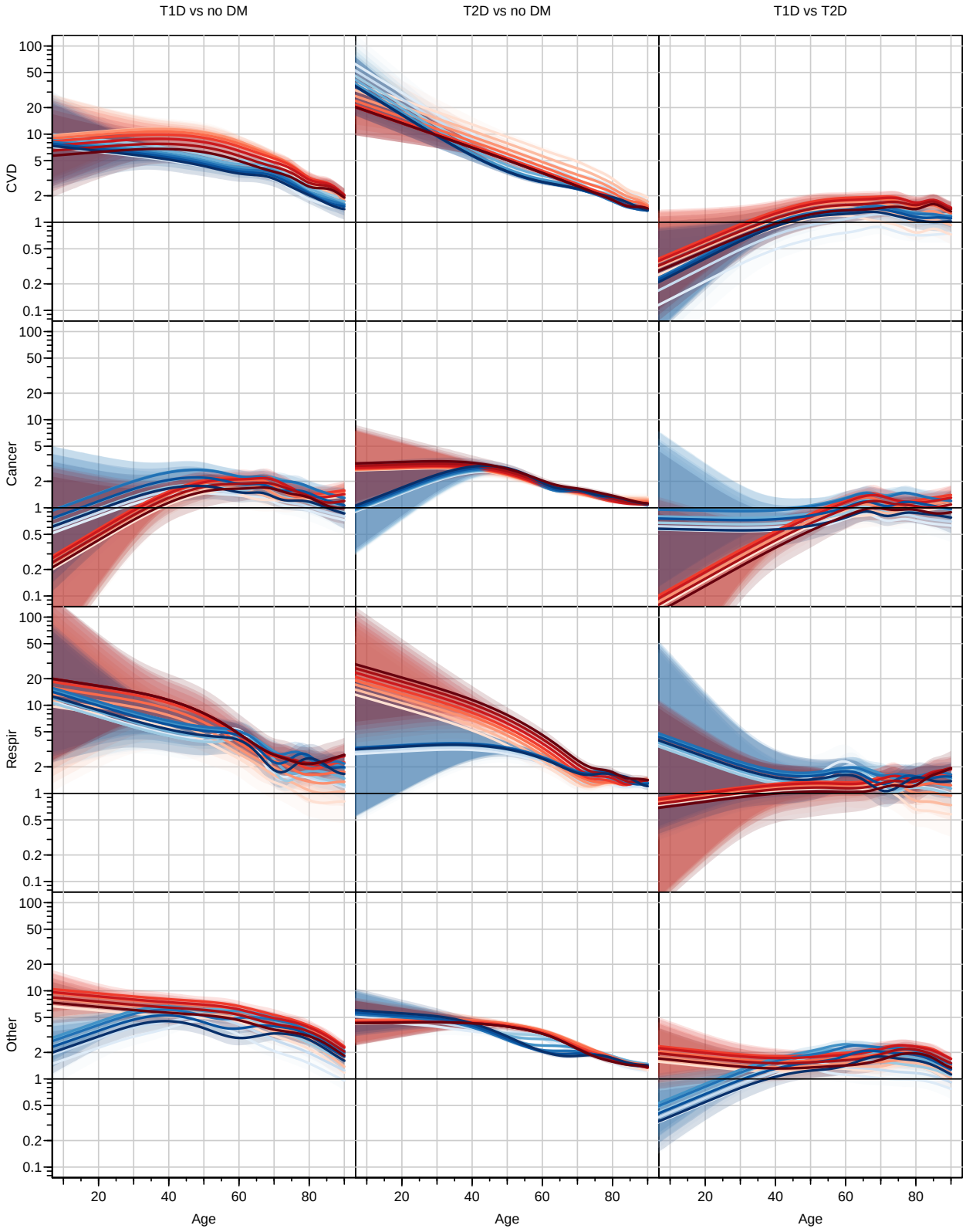


Figure 12.7: *Cause-specific mortality rate ratios between T1, T2 and noDM, based on cross sectional rates for the dates 1996, 1999,. . . ,2017, colouring from light to dark. Blue is for men, red for women; shading indicates 95% c.i.* `./graph/yll-cod-ratios`

```
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ age : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ from: chr [1:7] "noDM" "T1D" "T2D" "D-CVD" ...
  ..$ to  : chr [1:7] "noDM" "T1D" "T2D" "D-CVD" ...
```

So we now fill the transition probabilities in; note that the variable `Y` in the dataset `rt` is measured in millenia:

```
> round( addmargins( xtabs( Y ~ floor(P)+sex, data=rt ), 2 ), 1 )[-(1:15),]
        sex
floor(P)      M      W    Sum
    2011 2768.8 2811.4 5580.2
    2012 2780.9 2821.0 5601.9
    2013 2793.8 2831.9 5625.7
    2014 2811.7 2846.2 5657.9
    2015 2836.5 2864.2 5700.7
    2016 2854.0 2876.5 5730.5
```

Hence, when we make predictions for 1-month intervals we must do this for a value of Y equal to 1/12000 (that is `int`/1000).

We fill in the transition probabilities by just using the cumulative intensities over the intervals; when these are so small as in this case they are extremely close to the transition probabilities:

```
> for( sx in dimnames(Tr)[[1]] )
+ for( dd in dimnames(Tr)[[2]] )
+    {
+    nd <- data.frame( Px = as.numeric(dd),
+                      Ax = a.pt,
+                       Y = int/1000 )
+    Tr[sx,dd,,"noDM","T1D"]   <- ci.pred(  nDT1[[sx]], nd )[,1]
+    Tr[sx,dd,,"noDM","T2D"]   <- ci.pred(  nDT2[[sx]], nd )[,1]
+
+    Tr[sx,dd,,"noDM","D-CVD"] <- ci.pred( nDcvd[[sx]], nd )[,1]
+    Tr[sx,dd,,"noDM","D-Can"] <- ci.pred( nDcan[[sx]], nd )[,1]
+    Tr[sx,dd,,"noDM","D-Res"] <- ci.pred( nDres[[sx]], nd )[,1]
+    Tr[sx,dd,,"noDM","D-Oth"] <- ci.pred( nDoth[[sx]], nd )[,1]
+
+    Tr[sx,dd,, "T1D","D-CVD"] <- ci.pred( T1cvd[[sx]], nd )[,1]
+    Tr[sx,dd,, "T1D","D-Can"] <- ci.pred( T1can[[sx]], nd )[,1]
+    Tr[sx,dd,, "T1D","D-Res"] <- ci.pred( T1res[[sx]], nd )[,1]
+    Tr[sx,dd,, "T1D","D-Oth"] <- ci.pred( T1oth[[sx]], nd )[,1]
+
+    Tr[sx,dd,, "T2D","D-CVD"] <- ci.pred( T2cvd[[sx]], nd )[,1]
+    Tr[sx,dd,, "T2D","D-Can"] <- ci.pred( T2can[[sx]], nd )[,1]
+    Tr[sx,dd,, "T2D","D-Res"] <- ci.pred( T2res[[sx]], nd )[,1]
+    Tr[sx,dd,, "T2D","D-Oth"] <- ci.pred( T2oth[[sx]], nd )[,1]
+    }
> Tdiag <- 1 - apply( Tr, 1:4, sum )
> for( i in 1:dim(Tr)[5] ) Tr[,,,i,i] <- Tdiag[,,,i]
> range( Tr )
[1] 0 1
```

We can inspect three instances of the transition probability matrices; we multiply by 100,000 to get readable numbers (remember these are 1-month transition probabilities).

```
> round( ftable(Tr["M","2016",c(30,50,70)*12,,]*10^5, row.vars=1:2 ) )
                        to    noDM    T1D    T2D  D-CVD  D-Can  D-Res  D-Oth
age              from
29.9583333333333 noDM       99990      2      4      0      1      0      3
                 T1D            0  99984      0      2      1      1     13
                 T2D            0      0  99980      3      1      0     15
                 D-CVD          0      0      0 100000      0      0      0
                 D-Can          0      0      0      0 100000      0      0
                 D-Res          0      0      0      0      0 100000      0
                 D-Oth          0      0      0      0      0      0 100000
49.9583333333333 noDM       99934      1     43      4      6      1     11
                 T1D            0  99923      0     16     11      5     45
                 T2D            0      0  99933     14     16      3     34
                 D-CVD          0      0      0 100000      0      0      0
                 D-Can          0      0      0      0 100000      0      0
                 D-Res          0      0      0      0      0 100000      0
                 D-Oth          0      0      0      0      0      0 100000
69.9583333333333 noDM       99753      0     96     31     67     15     38
                 T1D            0  99643      0     98     96     29    134
                 T2D            0      0  99724     74    106     25     71
                 D-CVD          0      0      0 100000      0      0      0
                 D-Can          0      0      0      0 100000      0      0
                 D-Res          0      0      0      0      0 100000      0
                 D-Oth          0      0      0      0      0      0 100000
```

12.4 State occupancy probabilities

We can now compute state occupancy probabilities from various starting points; we have three starting states: noDM, T1D and T2D and starting ages for calculation of the residual life time. Note that we are computing the state occupancy probabilities at the borders of the age-intervals, using starting ages (conditioning on being in one of the transient states, that is alive) from 0 to 90 by 2 years.

```
> Pr <- ZArray( list( sex = dimnames(Tr)[[1]],
+                     date = 1996:2017,
+                 state.in = states[1:3],
+                   age.in = 0:99,
+                      age = c(0,a.pt+int/2), # border of age intervals
+                    state = states ) )
> str( Pr )
> fC( length( Pr ) )
> fC( prod( dim(Pr)[-5]) )
```

. . . so it's a pretty large object. Now it only remains to fill values into the object:

```
> system.time(
+ for( sx in dimnames(Pr)[[1]] ) # sex
+ for( dd in dimnames(Pr)[[2]] ) # date.in
+ for( st in dimnames(Pr)[[3]] ) # state.in
+ for( ii in dimnames(Pr)[[4]] ) # age.in
+    {
+    # ages after age.i
+    aa <- as.numeric(dimnames(Pr)[[5]]) > as.numeric(ii)
+    # for ages before the probability is 1 for the initial state
```

```
+    Pr[sx,dd,st,ii,!aa,st] <- 1
+    # and transitions occur at the remaining ages
+    for( ca in which(aa) )
+       Pr[sx,dd,st,ii,ca,] <- Pr[sx,dd,st,ii,ca-1,] %*% Tr[sx,dd,ca-1,,]
+    } )
> str( Pr )
> save( Pr, file="../nydata/Pr.Rda" ) # 344 Mb!


> system.time( load( file="../nydata/Pr.Rda" ) )
   user  system elapsed
   4.64    0.34    8.05
> # check sanity of resulting array:
> range( Pr )
[1] 0 1
> range( apply( Pr, 1:5, sum ) )
[1] 1 1
```

Thus we now have the state occupancy probabilities for all 6 states by age. And each of these classified by the conditioning variables (entry characteristics): sex, disease status, age and calendar time at entry)

12.5 State probabilities for 2016

We can now plot the cumulative risks of each of the 4 causes of death for select conditioning ages, 20, 40, 60 and 70; in each panel we will plot the cumulative risks fro T1, T2 and noDM for men and women together.

```
> a.fin <- 100
> a.ini <- c(40,50,60,70)
> ( d.st <- dimnames(Pr)[["state"]][4:7] )
[1] "D-CVD" "D-Can" "D-Res" "D-Oth"
> a1 <- function(){
+ axis( side=1, at=seq(0,90,10) )
+ axis( side=1, at=seq(0,95, 5), labels=NA, tcl=-0.2 )
+ }
> a2 <-
+ function()
+ {
+ axis(side=2,at=0:5/10,labels=c('',paste(1:4/10),'') )
+ axis(side=2,at=1:10/20,tcl=-0.3,labels=NA)
+ }
> par( mar=c(0,0,0,0), oma=c(4,4,2,1), mfrow=c(4,4),
+     las=1, bty="o" )
> layout( matrix(1:16,4,4,byrow=TRUE), widths=a.fin-a.ini )
> i <- 0
> for( st in d.st )
+ for( ai in a.ini )
+    {
+    plot( NA, xlim=c(ai,a.fin), ylim=c(0,0.42),
+              xaxt="n",         yaxt="n",
+              xaxs="i",         yaxs="i",
+              xlab="",          ylab="" )
```

```
+    abline(h=1:9/20,col=gray(0.8))
+    matlines( as.numeric(dimnames(Pr)[[5]]),
+              cbind( t(Pr["M","2016",,paste(ai),,st]),
+                     t(Pr["W","2016",,paste(ai),,st]) ),
+              col=rep(c("blue","red"),each=3),
+              lty=c("solid","11","31"), lend="butt", lwd=2 )
+    if( ai==a.ini[1] ) a2()
+    if( st==d.st[4] )  a1()
+    if( ai==a.ini[1] & st==d.st[1] ){
+        segments( 42, txy<-(17.5+0:2*5)/100, 52, txy,
+                  lty=c("solid","11","31"), lend="butt", lwd=2 )
+        text( 55, txy, dimnames(Pr)[[3]], adj=c(0,0.5) )
+        }
+    i <- i+1 ; text( ai+2, 0.375, letters[i], adj=0 )
+    }
> mtext( c("Other","Respir","Cancer","CVD"),
+        side=2, at=c(1,3,5,7)/8, line=3, cex=0.66, outer=T, las=0 )
> mtext( "Age", side=1, line=2, cex=0.66 )
```

12.5.1 State probabilities from age 50

We can plot the state occupancy probabilities for T1D resp T2D men aged 50, using the mortality rates as of 2016. Note that for persons starting in the no DM state, we need to merge the T1 and T2 states into the alive state; in the first instance we are not interested time spent with diabetes, only time spent alive.

```
> pr1 <- Pr["M","2016","T1D","50",,c(2,7:4)]
> pr2 <- Pr["M","2016","T2D","50",,c(3,7:4)]
> prn <- Pr["M","2016","noDM","50",,c(1:3,7:4)]
> # collapse the alive states
> prn[,1] <- prn[,1]+prn[,2]+prn[,3]
> prn <- prn[,-(2:3)]
> str( prn )
 num [1:1201, 1:5] 1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, "dimnames")=List of 2
  ..$ age  : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ state: chr [1:5] "noDM" "D-Oth" "D-Res" "D-Can" ...

> str( pr1 )
 num [1:1201, 1:5] 1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, "dimnames")=List of 2
  ..$ age  : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ state: chr [1:5] "T1D" "D-Oth" "D-Res" "D-Can" ...

> cr1 <- t( apply( cbind(0,pr1), 1, cumsum ))
> cr2 <- t( apply( cbind(0,pr2), 1, cumsum ))
> crn <- t( apply( cbind(0,prn), 1, cumsum ))
> str( crn )
 num [1:1201, 1:6] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 2
  ..$ : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ : chr [1:6] "" "noDM" "D-Oth" "D-Res" ...

> str( cr1 )
```

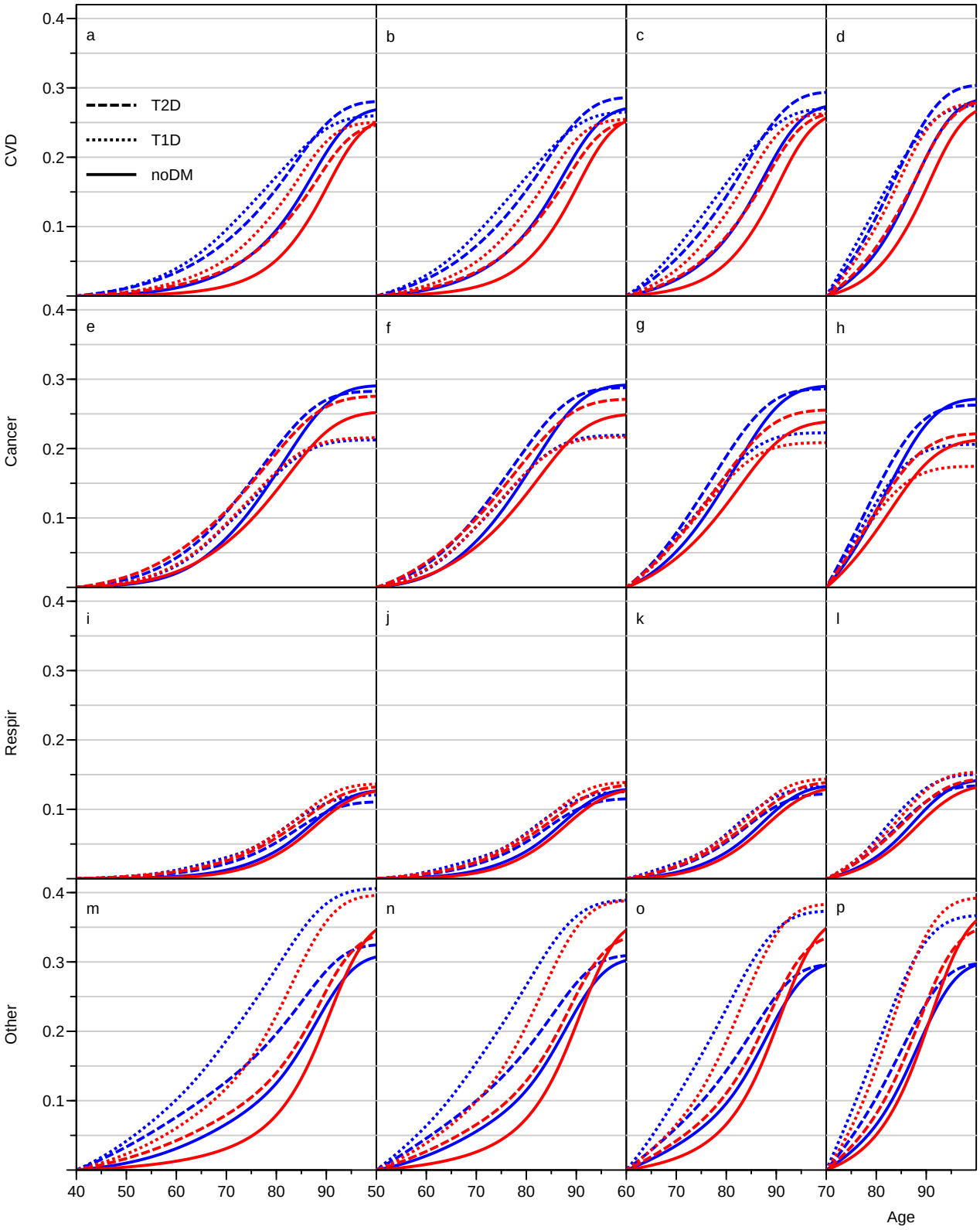


Figure 12.8: *Cumulative risk from ages, 40, 50, 60 and 70 for the 4 different causes of death. Solid lines are persons starting without DM, dotted lines T1D and broken lines T2D, red lines are men, blue lines are women.* `./graph/yll-cod-cumrisk`

```
 num [1:1201, 1:6] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 2
  ..$ : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ : chr [1:6] "" "T1D" "D-Oth" "D-Res" ...
> pr1 <- Pr["W","2016","T1D","50",,c(2,7:4)]
> pr2 <- Pr["W","2016","T2D","50",,c(3,7:4)]
> prn <- Pr["W","2016","noDM","50",,c(1:3,7:4)]
> # collapse the alive states
> prn[,1] <- prn[,1]+prn[,2]+prn[,3]
> prn <- prn[,-(2:3)]
> str( prn )
 num [1:1201, 1:5] 1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, "dimnames")=List of 2
  ..$ age  : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ state: chr [1:5] "noDM" "D-Oth" "D-Res" "D-Can" ...
> str( pr1 )
 num [1:1201, 1:5] 1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, "dimnames")=List of 2
  ..$ age  : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ state: chr [1:5] "T1D" "D-Oth" "D-Res" "D-Can" ...
> cr1 <- t( apply( cbind(0,pr1), 1, cumsum ))
> cr2 <- t( apply( cbind(0,pr2), 1, cumsum ))
> crn <- t( apply( cbind(0,prn), 1, cumsum ))
> str( crn )
 num [1:1201, 1:6] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 2
  ..$ : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ : chr [1:6] "" "noDM" "D-Oth" "D-Res" ...
> str( cr1 )
 num [1:1201, 1:6] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 2
  ..$ : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ : chr [1:6] "" "T1D" "D-Oth" "D-Res" ...
> filc <-
+ function( x, m, col=gray(0.6) )
+ polygon( c(x,rev(x)), c(m[,1],rev(m[,2])),
+          col=col,border="transparent" )
> clr <- brewer.pal(5,"Reds" )[-1]
> clr <- c("white",clr[1:4])
> par( mfrow=c(2,2), mar=c(3,0,1,0), oma=c(0,3,0,3), mgp=c(3,1,0)/1.6 )
> matplot( aa <- as.numeric(dimnames(cr1)[[1]]),
+          cr1,
+          xlim=c(50,100), ylim=0:1,
+          xaxs="i", yaxs="i",
+          lty=1, type="l",
+          xlab="Age at follow-up")
> for( i in 1:5 ) filc( aa, cr1[,i+1:0], clr[i] )
> text( 55, 0.1, "T1D", font=2 )
> box()
> matplot( aa <- as.numeric(dimnames(crn)[[1]]),
+          crn,
+          xlim=c(100,50), ylim=0:1,
+          xaxs="i", yaxs="i", yaxt="n",
+          lty=1, type="l",
```

```
+          xlab="Age at follow-up")
> for( i in 1:5 ) filc( aa, crn[,i+1:0], clr[i] )
> text( 55, 0.1, "no DM", font=2 )
> box()
> axis(side=4)
> ept <- crn[dim(crn)[1],]
> yy <- ept[3:6] - diff(ept[-1])/5
> text( 99, yy, names(ept)[3:6], col=c("black","white")[c(1,1,2,2)], font=2, adj=0, cex=1.2 )
> matplot( aa <- as.numeric(dimnames(cr2)[[1]]),
+          cr2,
+          xlim=c(50,100), ylim=0:1,
+          xaxs="i", yaxs="i",
+          lty=1, type="l",
+          xlab="Age at follow-up")
> for( i in 1:5 ) filc( aa, cr2[,i+1:0], clr[i] )
> text( 55, 0.1, "T2D", font=2 )
> box()
> matplot( aa <- as.numeric(dimnames(crn)[[1]]),
+          crn,
+          xlim=c(100,50), ylim=0:1,
+          xaxs="i", yaxs="i", yaxt='n',
+          lty=1, type="l",
+          xlab="Age at follow-up")
> for( i in 1:5 ) filc( aa, crn[,i+1:0], clr[i] )
> text( 55, 0.1, "no DM", font=2 )
> box()
> axis(side=4)
> ept <- crn[dim(crn)[1],]
> yy <- ept[3:6] - diff(ept[-1])/5
> text( 99, yy, names(ept)[3:6], col=c("black","white")[c(1,1,2,2)], font=2, adj=0, cex=1.2 )
```

12.6 Lifetime lost

The life-time lost from a given age, a, is the difference between the expected sojourn times in the alive states (no DM, T1D and T2D) for persons that are in state no DM at age a and persons in states T1D and T2D at age a. This is also (minus) the difference between the sojourn time in the four dead states, so the difference can be split by cause of death. These are what we shall call lifetime lost to different causes of death.

The sojourn time is simply the integral of the state probabilities from age a to ∞ (in practice 100 years), but we are only interested on the sojourn time in the death states[1], and these were are conveniently set to 0 for the ages prior to the conditioning age, so the sojourn times are just the sum, multiplied by the interval length:

```
> system.time( Sj <- apply( Pr[,,,,,4:7], c(1:4,6), sum ) * int )
   user  system elapsed
   2.69    0.42    3.10
> str( Sj )
 num [1:2, 1:22, 1:3, 1:100, 1:4] 9.47 7.17 9.35 7.09 9.22 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex     : chr [1:2] "M" "W"
```

[1] This is slightly counterintuitive but the differences between the "not lived" time after each of the causes are the cause-specific lifetime lost, and the sum of these to total lifetime lost.

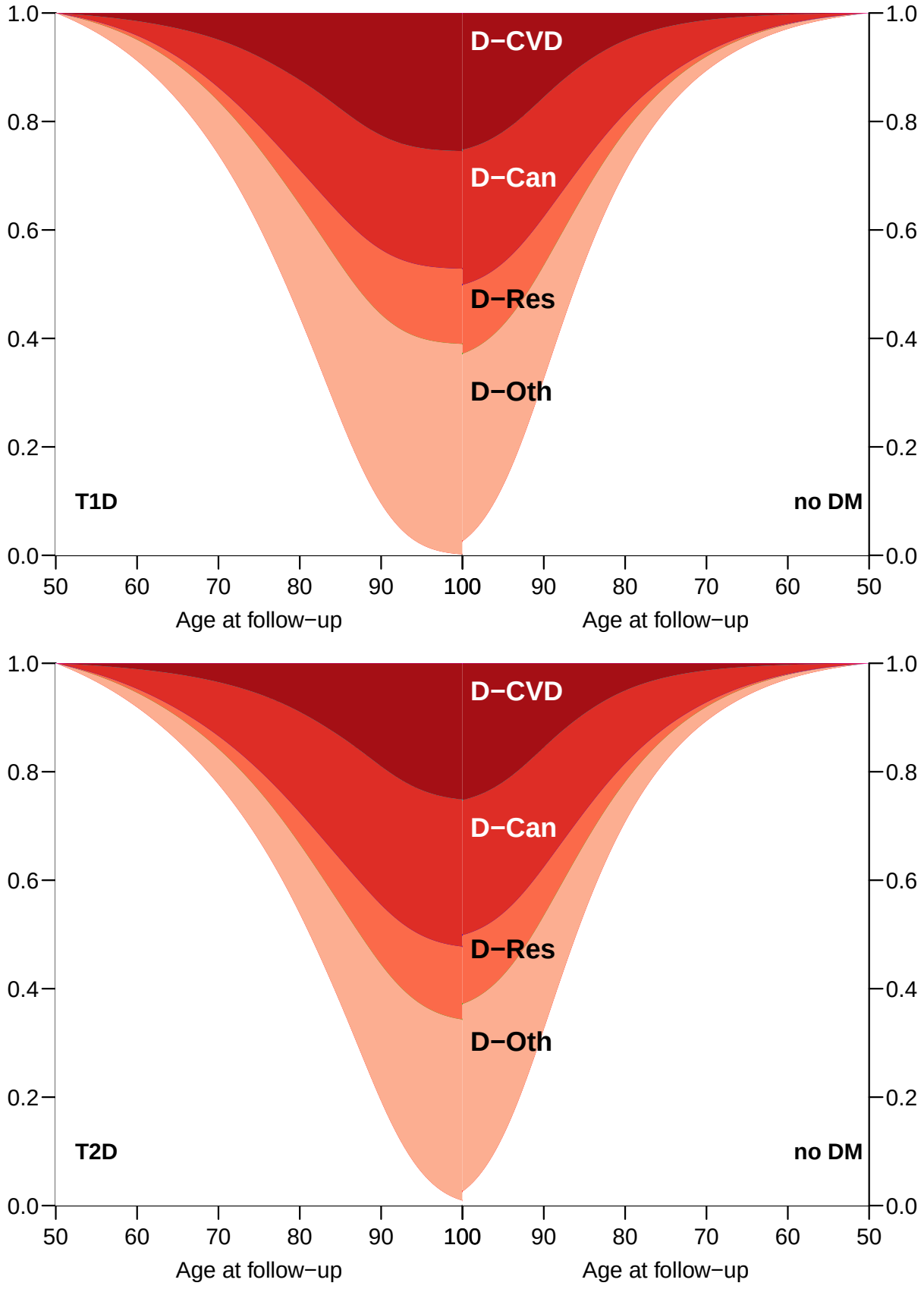


Figure 12.9: *Cumulative probabilities of death from three causes for men aged 50, with T1 resp. T2 diabetes (left) compared to persons without diabetes (right).* `./graph/yll-cod-prcau-M`

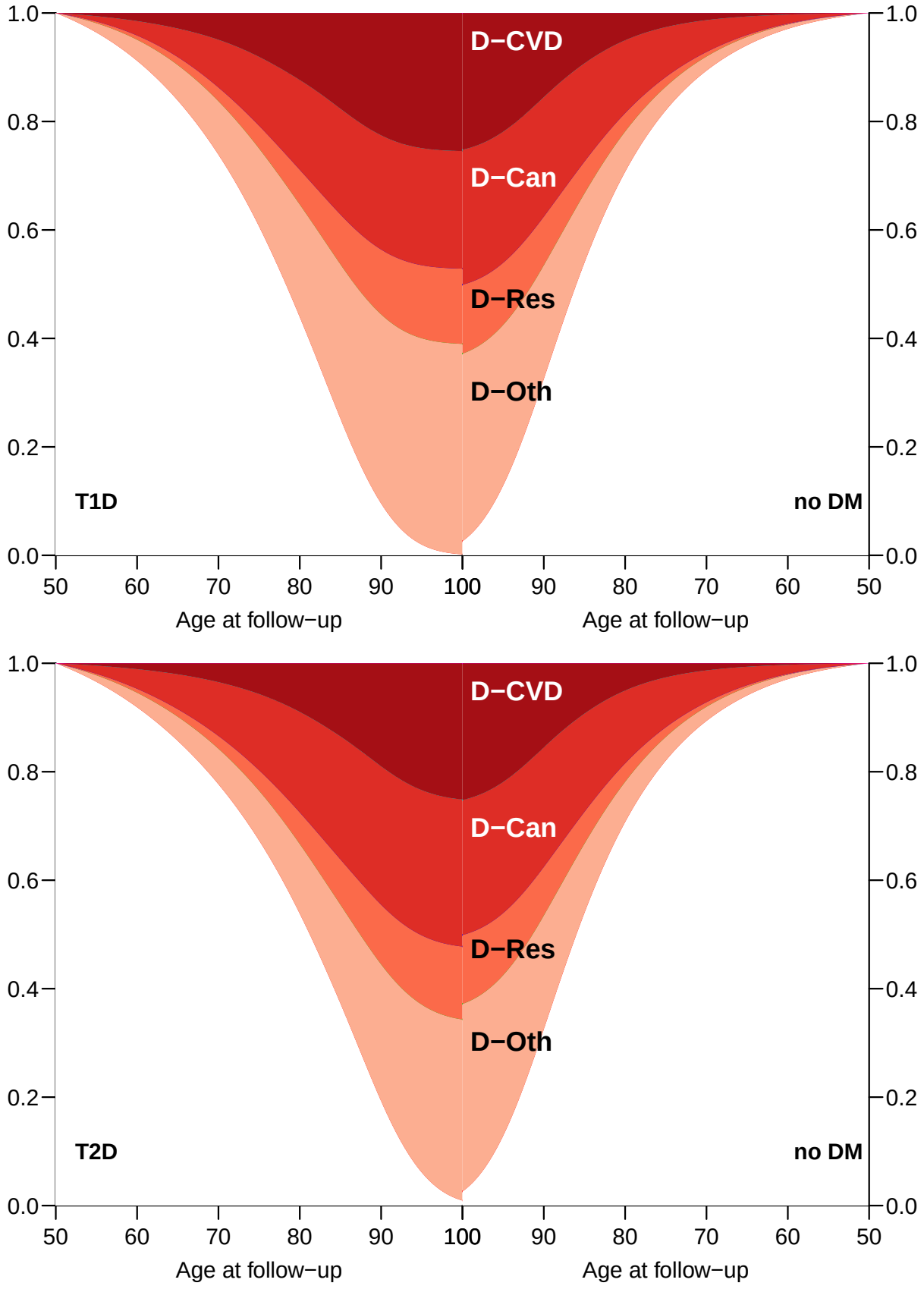


Figure 12.10: *Cumulative probabilities of death from three causes for women aged 50, with T1 resp. T2 diabetes (left) compared to8 persons without diabetes (right).* `./graph/yll-cod-prcau-W`

```
  ..$ date    : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ state.in: chr [1:3] "noDM" "T1D" "T2D"
  ..$ age.in  : chr [1:100] "0" "1" "2" "3" ...
  ..$ state   : chr [1:4] "D-CVD" "D-Can" "D-Res" "D-Oth"
```

The array `Sj` now contains the sojourn times in the four different death states (last dimension) for persons starting in any combination of the 4 conditions sex (`sex`), date (`date`), age (`age.in`) and state (`state.in`), where `date` refers to the date of the cross-sectional rates used.

So we want the YLL attributable to T1D resp. T2D but also subdivided by cause of death, so we need an array of the same structure as `Sj` except for the first level of the `state.in` dimension (`noDM`):

```
> YLL <- Sj[,,-1,,] * 0
> YLL[,,"T1D",,] <- Sj[,,"T1D",,] - Sj[,,"noDM",,]
> YLL[,,"T2D",,] <- Sj[,,"T2D",,] - Sj[,,"noDM",,]
> names( dimnames(YLL) )[3:5] <- c("S.in","A.in","CoD")
> YLL <- addmargins( YLL, 5 )
> str( YLL )
 num [1:2, 1:22, 1:2, 1:100, 1:5] 3.37 6.04 3.69 6.08 4 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
  ..$ A.in: chr [1:100] "0" "1" "2" "3" ...
  ..$ CoD : chr [1:5] "D-CVD" "D-Can" "D-Res" "D-Oth" ...
> save( YLL, file="../nydata/Yll.Rda" )
```

Thus we have the years of life lost to T1D resp. T2D diabetes at different index ages — note however that these are are not necessarily dates of *diagnosis* of diabetes — duration of disease is not considered here.

```
> load( file="../nydata/Yll.Rda" )
> round( ftable( YLL[,paste(seq(1996,2017,3)),,paste(c(25,4:6*10,75)),"Sum"],
+                col.vars=3:4, row.vars=1:2 ), 1 )
         S.in  T1D                      T2D
         A.in   25   40   50   60   75   25   40   50   60   75
sex date
M   1996       9.6  8.3  6.6  4.7  1.8 11.4  9.5  7.7  5.7  2.9
    1999      11.4 10.0  8.2  6.1  2.7 10.6  8.7  7.1  5.2  2.7
    2002      12.7 11.1  9.3  6.9  3.3  9.8  8.1  6.5  4.9  2.4
    2005      13.2 11.6  9.7  7.3  3.6  9.0  7.3  5.9  4.4  2.2
    2008      13.6 12.0 10.1  7.8  4.0  7.8  6.4  5.1  3.9  2.0
    2011      12.4 11.0  9.4  7.5  4.0  7.0  5.7  4.6  3.5  1.9
    2014      10.2  9.1  7.9  6.4  3.6  6.9  5.6  4.6  3.6  2.0
    2017       8.1  7.3  6.3  5.2  3.0  6.9  5.7  4.7  3.7  2.1
W   1996      10.9  9.7  8.3  6.4  3.0 11.0  9.5  8.1  6.3  3.5
    1999      11.5 10.3  9.1  7.1  3.7  9.9  8.6  7.3  5.7  3.2
    2002      11.9 10.8  9.5  7.7  4.2  9.0  7.8  6.6  5.1  2.9
    2005      12.4 11.2  9.9  8.1  4.6  8.0  7.0  5.9  4.6  2.5
    2008      12.6 11.4 10.2  8.4  5.0  7.1  6.2  5.2  4.0  2.1
    2011      11.4 10.4  9.4  7.9  4.8  6.5  5.7  4.8  3.7  2.0
    2014       9.8  9.0  8.1  6.9  4.4  6.4  5.6  4.8  3.7  2.0
    2017       8.2  7.6  6.9  6.0  4.0  6.4  5.7  4.9  3.8  2.1
> round( ftable( YLL[,paste(seq(1996,2017,3)),,paste(2:8*10),],
+                col.vars=c(5,1), row.vars=c(3,2,4) ), 1 )
```

		CoD	D-CVD		D-Can		D-Res		D-Oth		Sum	
		sex	M	W	M	W	M	W	M	W	M	W
S.in	date	A.in										
T1D	1996	20	3.4	6.2	-2.0	-1.3	0.4	-0.1	7.9	6.4	9.7	11.2
		30	3.5	6.2	-2.0	-1.1	0.4	-0.1	7.3	5.7	9.3	10.6
		40	3.7	6.1	-1.9	-1.0	0.4	-0.1	6.1	4.7	8.3	9.7
		50	4.0	5.7	-1.6	-0.8	0.5	-0.2	3.6	3.6	6.6	8.3
		60	4.2	5.0	-1.3	-0.7	0.4	-0.4	1.4	2.3	4.7	6.4
		70	3.4	3.7	-1.2	-0.4	0.1	-0.5	0.4	1.3	2.7	4.1
		80	1.7	1.8	-0.8	0.0	0.2	-0.4	-0.1	0.6	1.1	2.0
	1999	20	4.4	6.2	-2.2	-1.0	0.0	0.0	9.4	6.5	11.6	11.8
		30	4.5	6.3	-2.1	-0.9	0.0	0.0	8.7	5.8	11.1	11.2
		40	4.6	6.2	-2.0	-0.7	0.0	0.0	7.4	4.9	10.0	10.3
		50	4.9	5.8	-1.7	-0.5	0.1	-0.1	4.9	3.8	8.2	9.1
		60	5.1	5.2	-1.4	-0.5	0.0	-0.2	2.3	2.6	6.1	7.1
		70	4.2	4.1	-1.2	-0.5	-0.2	-0.3	1.0	1.6	3.7	4.9
		80	2.3	2.2	-0.8	-0.1	0.0	-0.3	0.3	0.8	1.8	2.6
	2002	20	4.3	5.7	-2.1	-0.6	-0.3	0.2	11.0	6.9	12.9	12.2
		30	4.4	5.7	-2.0	-0.4	-0.3	0.2	10.2	6.1	12.3	11.6
		40	4.6	5.7	-1.9	-0.3	-0.3	0.1	8.7	5.2	11.1	10.8
		50	4.8	5.4	-1.6	-0.1	-0.2	0.1	6.3	4.1	9.3	9.5
		60	4.9	4.9	-1.2	-0.2	-0.2	0.0	3.4	3.0	6.9	7.7
		70	4.1	4.0	-1.1	-0.5	-0.4	-0.1	1.8	2.0	4.4	5.4
		80	2.3	2.3	-0.8	-0.1	-0.1	-0.2	0.9	1.1	2.3	3.1
	2005	20	2.5	4.2	-1.4	0.1	-0.2	0.3	12.6	8.0	13.5	12.7
		30	2.6	4.3	-1.3	0.2	-0.2	0.3	11.8	7.2	12.9	12.0
		40	2.7	4.2	-1.1	0.4	-0.2	0.3	10.2	6.2	11.6	11.2
		50	2.8	4.1	-0.8	0.6	-0.1	0.3	7.8	5.1	9.7	9.9
		60	3.0	3.7	-0.4	0.4	0.0	0.2	4.8	3.8	7.3	8.1
		70	2.4	3.2	-0.5	-0.2	-0.1	0.1	2.9	2.8	4.8	5.8
		80	1.3	1.9	-0.5	-0.1	0.1	0.0	1.6	1.7	2.6	3.5
	2008	20	1.3	3.0	-0.3	0.5	0.1	0.5	12.8	8.9	13.9	12.9
		30	1.3	3.0	-0.2	0.7	0.1	0.5	12.0	8.1	13.2	12.2
		40	1.4	3.0	0.0	0.9	0.1	0.5	10.4	7.0	12.0	11.4
		50	1.5	2.9	0.3	1.0	0.2	0.4	8.1	5.8	10.1	10.2
		60	1.5	2.7	0.6	0.9	0.2	0.4	5.5	4.5	7.8	8.4
		70	1.2	2.3	0.4	0.1	0.3	0.3	3.4	3.5	5.3	6.2
		80	0.5	1.5	0.0	-0.1	0.4	0.2	2.0	2.3	2.9	3.9
	2011	20	1.3	2.4	0.1	0.6	0.4	0.7	10.9	8.0	12.6	11.7
		30	1.3	2.5	0.2	0.7	0.4	0.7	10.2	7.3	12.0	11.1
		40	1.4	2.5	0.3	0.9	0.4	0.7	8.9	6.4	11.0	10.4
		50	1.4	2.4	0.6	1.0	0.4	0.6	7.0	5.3	9.4	9.4
		60	1.3	2.2	0.8	0.9	0.4	0.6	5.1	4.2	7.5	7.9
		70	0.9	1.9	0.5	0.1	0.5	0.5	3.3	3.3	5.2	5.9
		80	0.3	1.3	0.1	-0.2	0.6	0.5	1.9	2.2	2.8	3.8
	2014	20	1.8	2.2	-0.4	0.4	0.6	0.9	8.4	6.6	10.4	10.0
		30	1.8	2.2	-0.3	0.4	0.6	0.9	7.9	6.0	10.0	9.5
		40	1.8	2.2	-0.2	0.6	0.6	0.8	7.0	5.3	9.1	9.0
		50	1.8	2.1	0.0	0.7	0.6	0.8	5.6	4.5	7.9	8.1
		60	1.5	2.0	0.0	0.6	0.5	0.7	4.3	3.6	6.4	6.9
		70	1.0	1.7	-0.1	0.1	0.6	0.7	3.0	2.9	4.5	5.4
		80	0.4	1.2	-0.3	-0.3	0.7	0.7	1.8	1.9	2.5	3.5
	2017	20	2.1	1.9	-1.0	0.1	0.7	1.0	6.4	5.4	8.2	8.4
		30	2.1	1.9	-1.0	0.1	0.7	1.0	6.1	5.0	7.9	8.0
		40	2.1	1.9	-0.9	0.2	0.7	1.0	5.4	4.4	7.3	7.6
		50	2.0	1.9	-0.8	0.3	0.7	1.0	4.4	3.8	6.3	6.9
		60	1.7	1.7	-0.7	0.3	0.6	0.9	3.5	3.1	5.2	6.0

		70	1.2	1.5	-0.8	-0.1	0.6	0.8	2.7	2.5	3.7	4.7
		80	0.5	1.1	-0.6	-0.4	0.7	0.8	1.6	1.6	2.2	3.2
T2D	1996	20	6.8	6.6	-0.6	1.0	-0.7	-0.1	6.4	3.8	11.9	11.3
		30	6.8	6.4	-0.4	1.0	-0.6	-0.1	5.2	3.3	10.9	10.6
		40	6.6	6.1	-0.3	1.0	-0.6	-0.2	3.8	2.6	9.5	9.5
		50	6.3	5.9	-0.2	0.7	-0.6	-0.2	2.2	1.7	7.7	8.1
		60	5.8	5.6	-0.3	0.2	-0.6	-0.3	0.8	0.8	5.7	6.3
		70	4.7	4.9	-0.3	0.0	-0.5	-0.3	-0.1	-0.1	3.8	4.5
		80	2.9	3.0	-0.3	0.0	-0.3	-0.1	-0.3	-0.3	2.0	2.5
	1999	20	5.2	5.1	-0.3	1.1	-0.5	0.0	6.6	3.9	11.1	10.2
		30	5.2	4.9	-0.2	1.2	-0.5	0.0	5.5	3.5	10.0	9.6
		40	5.0	4.7	0.0	1.1	-0.4	0.0	4.1	2.8	8.7	8.6
		50	4.8	4.5	0.0	0.8	-0.4	0.0	2.7	2.0	7.1	7.3
		60	4.5	4.3	-0.1	0.4	-0.4	-0.1	1.3	1.2	5.2	5.7
		70	3.6	3.9	-0.2	0.0	-0.4	-0.2	0.4	0.3	3.4	4.0
		80	2.3	2.4	-0.2	0.0	-0.2	-0.1	0.0	0.0	1.9	2.3
	2002	20	4.0	3.8	-0.1	1.2	-0.3	0.2	6.8	4.0	10.3	9.3
		30	3.9	3.7	0.0	1.2	-0.3	0.2	5.7	3.6	9.3	8.7
		40	3.8	3.5	0.2	1.2	-0.3	0.2	4.3	3.0	8.1	7.8
		50	3.6	3.4	0.2	0.9	-0.2	0.1	2.9	2.2	6.5	6.6
		60	3.3	3.2	0.1	0.5	-0.2	0.0	1.7	1.4	4.9	5.1
		70	2.6	2.9	-0.1	0.1	-0.2	-0.1	0.8	0.7	3.2	3.6
		80	1.7	1.9	-0.2	0.0	-0.1	0.0	0.4	0.3	1.8	2.1
	2005	20	3.0	2.7	0.1	1.3	-0.2	0.3	6.4	3.9	9.4	8.3
		30	3.0	2.6	0.2	1.3	-0.1	0.3	5.4	3.5	8.5	7.8
		40	2.9	2.5	0.4	1.3	-0.1	0.3	4.1	2.9	7.3	7.0
		50	2.8	2.4	0.4	1.0	-0.1	0.3	2.8	2.3	5.9	5.9
		60	2.5	2.3	0.2	0.6	0.0	0.2	1.8	1.5	4.4	4.6
		70	2.0	2.0	0.0	0.2	0.0	0.1	1.0	0.9	3.0	3.2
		80	1.3	1.4	-0.2	0.0	0.0	0.0	0.5	0.5	1.7	1.9
	2008	20	2.3	1.8	0.3	1.4	0.0	0.5	5.6	3.6	8.2	7.3
		30	2.3	1.8	0.5	1.4	0.0	0.5	4.6	3.2	7.4	6.9
		40	2.2	1.7	0.6	1.4	0.1	0.4	3.5	2.7	6.4	6.2
		50	2.1	1.6	0.6	1.1	0.1	0.4	2.4	2.1	5.1	5.2
		60	1.9	1.6	0.4	0.7	0.1	0.3	1.5	1.5	3.9	4.0
		70	1.5	1.4	0.1	0.2	0.1	0.2	0.9	0.9	2.6	2.7
		80	0.9	1.0	-0.1	0.0	0.1	0.1	0.5	0.6	1.4	1.6
	2011	20	1.9	1.4	0.5	1.5	0.1	0.6	4.8	3.2	7.4	6.7
		30	1.9	1.4	0.6	1.5	0.1	0.6	4.0	2.8	6.7	6.3
		40	1.8	1.3	0.7	1.4	0.2	0.6	3.0	2.4	5.7	5.7
		50	1.7	1.3	0.7	1.2	0.2	0.5	2.0	1.9	4.6	4.8
		60	1.6	1.2	0.5	0.8	0.2	0.4	1.3	1.3	3.5	3.7
		70	1.2	1.1	0.2	0.3	0.2	0.3	0.8	0.8	2.5	2.5
		80	0.7	0.7	-0.1	0.0	0.2	0.2	0.5	0.6	1.3	1.5
	2014	20	1.9	1.3	0.7	1.8	0.2	0.7	4.4	2.8	7.2	6.6
		30	1.9	1.2	0.8	1.7	0.2	0.7	3.6	2.5	6.5	6.2
		40	1.8	1.2	0.9	1.7	0.2	0.7	2.7	2.1	5.6	5.6
		50	1.7	1.1	0.9	1.4	0.2	0.6	1.8	1.7	4.6	4.8
		60	1.5	1.1	0.6	0.9	0.2	0.6	1.1	1.2	3.6	3.7
		70	1.2	0.9	0.3	0.5	0.2	0.4	0.8	0.7	2.5	2.6
		80	0.7	0.6	0.0	0.0	0.2	0.3	0.5	0.6	1.4	1.5
	2017	20	1.9	1.2	1.0	2.1	0.2	0.8	4.1	2.5	7.2	6.5
		30	1.9	1.1	1.1	2.1	0.2	0.8	3.4	2.2	6.5	6.2
		40	1.8	1.1	1.1	2.0	0.2	0.8	2.6	1.9	5.7	5.7
		50	1.7	1.0	1.1	1.7	0.2	0.7	1.7	1.5	4.7	4.9
		60	1.6	1.0	0.8	1.2	0.2	0.7	1.1	1.0	3.7	3.8
		70	1.2	0.8	0.4	0.7	0.2	0.5	0.8	0.6	2.7	2.7

```
          80         0.8  0.6   0.0  0.2   0.2  0.3   0.6  0.5  1.6  1.6
> round( ftable( YLL[,paste(seq(1996,2017,1)),,paste(4:7*10),],
+                col.vars=c(5,1), row.vars=c(3,4,2) ), 1 )
               CoD D-CVD      D-Can      D-Res      D-Oth       Sum
               sex     M    W     M    W     M    W     M    W    M    W
S.in A.in date
T1D  40   1996       3.7  6.1  -1.9 -1.0   0.4 -0.1   6.1  4.7  8.3  9.7
          1997       4.0  6.1  -1.9 -0.9   0.3 -0.1   6.5  4.8  8.9  9.9
          1998       4.3  6.1  -1.9 -0.8   0.1 -0.1   6.9  4.8  9.4 10.1
          1999       4.6  6.2  -2.0 -0.7   0.0  0.0   7.4  4.9 10.0 10.3
          2000       4.8  6.1  -2.0 -0.6  -0.1  0.0   7.8  4.9 10.5 10.5
          2001       4.8  6.0  -2.0 -0.4  -0.2  0.1   8.2  5.0 10.9 10.7
          2002       4.6  5.7  -1.9 -0.3  -0.3  0.1   8.7  5.2 11.1 10.8
          2003       4.0  5.2  -1.7 -0.1  -0.3  0.2   9.3  5.5 11.3 10.8
          2004       3.4  4.7  -1.4  0.2  -0.2  0.3   9.8  5.8 11.4 11.0
          2005       2.7  4.2  -1.1  0.4  -0.2  0.3  10.2  6.2 11.6 11.2
          2006       2.1  3.8  -0.7  0.6  -0.1  0.4  10.5  6.6 11.8 11.3
          2007       1.7  3.3  -0.3  0.8   0.0  0.4  10.5  6.9 11.9 11.4
          2008       1.4  3.0   0.0  0.9   0.1  0.5  10.4  7.0 12.0 11.4
          2009       1.3  2.7   0.3  0.9   0.2  0.6  10.0  6.9 11.8 11.2
          2010       1.3  2.6   0.4  0.9   0.3  0.6   9.5  6.7 11.5 10.8
          2011       1.4  2.5   0.3  0.9   0.4  0.7   8.9  6.4 11.0 10.4
          2012       1.5  2.4   0.2  0.8   0.5  0.7   8.2  6.0 10.4  9.9
          2013       1.7  2.3   0.0  0.7   0.5  0.8   7.6  5.7  9.8  9.4
          2014       1.8  2.2  -0.2  0.6   0.6  0.8   7.0  5.3  9.1  9.0
          2015       1.9  2.1  -0.4  0.5   0.6  0.9   6.4  5.0  8.5  8.5
          2016       2.0  2.0  -0.7  0.3   0.6  0.9   5.9  4.7  7.9  8.0
          2017       2.1  1.9  -0.9  0.2   0.7  1.0   5.4  4.4  7.3  7.6
     50   1996       4.0  5.7  -1.6 -0.8   0.5 -0.2   3.6  3.6  6.6  8.3
          1997       4.4  5.8  -1.6 -0.7   0.4 -0.2   4.1  3.7  7.1  8.6
          1998       4.7  5.8  -1.7 -0.6   0.2 -0.1   4.5  3.7  7.7  8.8
          1999       4.9  5.8  -1.7 -0.5   0.1 -0.1   4.9  3.8  8.2  9.1
          2000       5.1  5.8  -1.7 -0.4  -0.1  0.0   5.4  3.9  8.6  9.2
          2001       5.1  5.7  -1.7 -0.3  -0.2  0.0   5.8  4.0  9.0  9.4
          2002       4.8  5.4  -1.6 -0.1  -0.2  0.1   6.3  4.1  9.3  9.5
          2003       4.2  5.0  -1.4  0.1  -0.2  0.1   6.9  4.4  9.4  9.6
          2004       3.5  4.5  -1.2  0.3  -0.2  0.2   7.4  4.7  9.6  9.8
          2005       2.8  4.1  -0.8  0.6  -0.1  0.3   7.8  5.1  9.7  9.9
          2006       2.2  3.6  -0.4  0.8   0.0  0.3   8.1  5.4  9.9 10.1
          2007       1.8  3.2   0.0  0.9   0.1  0.4   8.2  5.7 10.1 10.2
          2008       1.5  2.9   0.3  1.0   0.2  0.4   8.1  5.8 10.1 10.2
          2009       1.3  2.7   0.6  1.1   0.3  0.5   7.9  5.8 10.1 10.0
          2010       1.3  2.5   0.7  1.1   0.4  0.6   7.5  5.6  9.8  9.7
          2011       1.4  2.4   0.6  1.0   0.4  0.6   7.0  5.3  9.4  9.4
          2012       1.5  2.3   0.4  0.9   0.5  0.7   6.5  5.1  9.0  9.0
          2013       1.6  2.2   0.2  0.8   0.5  0.8   6.1  4.8  8.4  8.6
          2014       1.8  2.1   0.0  0.7   0.6  0.8   5.6  4.5  7.9  8.1
          2015       1.9  2.0  -0.3  0.6   0.6  0.9   5.2  4.3  7.4  7.7
          2016       2.0  2.0  -0.5  0.5   0.7  0.9   4.8  4.0  6.8  7.3
          2017       2.0  1.9  -0.8  0.3   0.7  1.0   4.4  3.8  6.3  6.9
     60   1996       4.2  5.0  -1.3 -0.7   0.4 -0.4   1.4  2.3  4.7  6.4
          1997       4.5  5.1  -1.3 -0.6   0.3 -0.3   1.7  2.4  5.2  6.6
          1998       4.8  5.2  -1.4 -0.6   0.2 -0.2   2.0  2.5  5.6  6.9
          1999       5.1  5.2  -1.4 -0.5   0.0 -0.2   2.3  2.6  6.1  7.1
          2000       5.2  5.2  -1.4 -0.4  -0.1 -0.1   2.6  2.7  6.4  7.3
          2001       5.2  5.1  -1.3 -0.3  -0.1 -0.1   3.0  2.8  6.7  7.5
          2002       4.9  4.9  -1.2 -0.2  -0.2  0.0   3.4  3.0  6.9  7.7
```

		2003	4.4	4.6	-1.0	0.0	-0.2	0.0	3.9	3.2	7.1	7.8
		2004	3.7	4.2	-0.8	0.2	-0.1	0.1	4.4	3.5	7.2	8.0
		2005	3.0	3.7	-0.4	0.4	0.0	0.2	4.8	3.8	7.3	8.1
		2006	2.4	3.3	-0.1	0.6	0.0	0.2	5.2	4.1	7.5	8.3
		2007	1.9	3.0	0.3	0.8	0.1	0.3	5.4	4.4	7.7	8.4
		2008	1.5	2.7	0.6	0.9	0.2	0.4	5.5	4.5	7.8	8.4
		2009	1.3	2.5	0.8	0.9	0.3	0.4	5.4	4.5	7.9	8.3
		2010	1.2	2.3	0.9	0.9	0.4	0.5	5.3	4.4	7.7	8.1
		2011	1.3	2.2	0.8	0.9	0.4	0.6	5.1	4.2	7.5	7.9
		2012	1.3	2.1	0.6	0.8	0.5	0.6	4.8	4.0	7.2	7.6
		2013	1.4	2.1	0.3	0.7	0.5	0.7	4.6	3.8	6.8	7.3
		2014	1.5	2.0	0.0	0.6	0.5	0.7	4.3	3.6	6.4	6.9
		2015	1.6	1.9	-0.2	0.5	0.6	0.8	4.0	3.4	6.0	6.6
		2016	1.7	1.8	-0.5	0.4	0.6	0.8	3.8	3.3	5.6	6.3
		2017	1.7	1.7	-0.7	0.3	0.6	0.9	3.5	3.1	5.2	6.0
	70	1996	3.4	3.7	-1.2	-0.4	0.1	-0.5	0.4	1.3	2.7	4.1
		1997	3.7	3.9	-1.2	-0.4	0.0	-0.5	0.6	1.4	3.1	4.4
		1998	4.0	4.0	-1.2	-0.5	-0.1	-0.4	0.8	1.5	3.4	4.6
		1999	4.2	4.1	-1.2	-0.5	-0.2	-0.3	1.0	1.6	3.7	4.9
		2000	4.4	4.2	-1.2	-0.5	-0.3	-0.3	1.2	1.7	4.0	5.1
		2001	4.3	4.2	-1.2	-0.5	-0.4	-0.2	1.5	1.8	4.3	5.2
		2002	4.1	4.0	-1.1	-0.5	-0.4	-0.1	1.8	2.0	4.4	5.4
		2003	3.6	3.8	-0.9	-0.4	-0.3	-0.1	2.2	2.2	4.5	5.5
		2004	3.1	3.5	-0.7	-0.3	-0.3	0.0	2.6	2.5	4.6	5.7
		2005	2.4	3.2	-0.5	-0.2	-0.1	0.1	2.9	2.8	4.8	5.8
		2006	1.9	2.8	-0.1	-0.1	0.0	0.2	3.2	3.1	4.9	6.0
		2007	1.5	2.6	0.2	0.0	0.1	0.3	3.4	3.3	5.1	6.1
		2008	1.2	2.3	0.4	0.1	0.3	0.3	3.4	3.5	5.3	6.2
		2009	1.0	2.1	0.6	0.1	0.4	0.4	3.4	3.5	5.3	6.2
		2010	0.9	2.0	0.6	0.1	0.5	0.5	3.4	3.4	5.3	6.1
		2011	0.9	1.9	0.5	0.1	0.5	0.5	3.3	3.3	5.2	5.9
		2012	0.9	1.8	0.3	0.1	0.6	0.6	3.2	3.2	5.0	5.7
		2013	1.0	1.8	0.1	0.1	0.6	0.6	3.1	3.0	4.8	5.6
		2014	1.0	1.7	-0.1	0.1	0.6	0.7	3.0	2.9	4.5	5.4
		2015	1.1	1.6	-0.4	0.0	0.6	0.7	2.9	2.8	4.3	5.1
		2016	1.1	1.6	-0.6	0.0	0.6	0.8	2.8	2.6	4.0	4.9
		2017	1.2	1.5	-0.8	-0.1	0.6	0.8	2.7	2.5	3.7	4.7
T2D	40	1996	6.6	6.1	-0.3	1.0	-0.6	-0.2	3.8	2.6	9.5	9.5
		1997	6.0	5.7	-0.2	1.0	-0.6	-0.1	3.9	2.7	9.2	9.2
		1998	5.5	5.2	-0.1	1.1	-0.5	0.0	4.0	2.7	9.0	8.9
		1999	5.0	4.7	0.0	1.1	-0.4	0.0	4.1	2.8	8.7	8.6
		2000	4.6	4.3	0.0	1.1	-0.4	0.1	4.2	2.9	8.5	8.4
		2001	4.2	3.9	0.1	1.2	-0.3	0.1	4.3	2.9	8.3	8.1
		2002	3.8	3.5	0.2	1.2	-0.3	0.2	4.3	3.0	8.1	7.8
		2003	3.5	3.1	0.2	1.2	-0.2	0.2	4.3	3.0	7.8	7.6
		2004	3.2	2.8	0.3	1.2	-0.2	0.3	4.3	3.0	7.6	7.3
		2005	2.9	2.5	0.4	1.3	-0.1	0.3	4.1	2.9	7.3	7.0
		2006	2.7	2.2	0.4	1.3	-0.1	0.3	4.0	2.9	7.0	6.7
		2007	2.4	1.9	0.5	1.3	0.0	0.4	3.7	2.8	6.7	6.4
		2008	2.2	1.7	0.6	1.4	0.1	0.4	3.5	2.7	6.4	6.2
		2009	2.0	1.5	0.6	1.4	0.1	0.5	3.3	2.6	6.1	6.0
		2010	1.9	1.4	0.7	1.4	0.1	0.5	3.1	2.5	5.9	5.8
		2011	1.8	1.3	0.7	1.4	0.2	0.6	3.0	2.4	5.7	5.7
		2012	1.8	1.3	0.8	1.5	0.2	0.6	2.9	2.3	5.6	5.6
		2013	1.8	1.2	0.8	1.6	0.2	0.6	2.8	2.2	5.6	5.6
		2014	1.8	1.2	0.9	1.7	0.2	0.7	2.7	2.1	5.6	5.6
		2015	1.8	1.1	1.0	1.8	0.2	0.7	2.7	2.0	5.7	5.6

	2016	1.8	1.1	1.1	1.9	0.2	0.7	2.6	2.0	5.7	5.7
	2017	1.8	1.1	1.1	2.0	0.2	0.8	2.6	1.9	5.7	5.7
50	1996	6.3	5.9	-0.2	0.7	-0.6	-0.2	2.2	1.7	7.7	8.1
	1997	5.8	5.4	-0.1	0.8	-0.5	-0.1	2.4	1.8	7.5	7.8
	1998	5.3	4.9	0.0	0.8	-0.5	-0.1	2.5	1.9	7.3	7.6
	1999	4.8	4.5	0.0	0.8	-0.4	0.0	2.7	2.0	7.1	7.3
	2000	4.4	4.1	0.1	0.9	-0.3	0.0	2.8	2.1	6.9	7.1
	2001	4.0	3.7	0.2	0.9	-0.3	0.1	2.9	2.2	6.7	6.9
	2002	3.6	3.4	0.2	0.9	-0.2	0.1	2.9	2.2	6.5	6.6
	2003	3.3	3.0	0.3	1.0	-0.2	0.2	3.0	2.3	6.4	6.4
	2004	3.0	2.7	0.3	1.0	-0.1	0.2	2.9	2.3	6.2	6.2
	2005	2.8	2.4	0.4	1.0	-0.1	0.3	2.8	2.3	5.9	5.9
	2006	2.5	2.1	0.5	1.1	0.0	0.3	2.7	2.2	5.7	5.7
	2007	2.3	1.8	0.5	1.1	0.0	0.3	2.6	2.2	5.4	5.4
	2008	2.1	1.6	0.6	1.1	0.1	0.4	2.4	2.1	5.1	5.2
	2009	1.9	1.5	0.6	1.1	0.1	0.4	2.2	2.0	4.9	5.0
	2010	1.8	1.3	0.7	1.2	0.2	0.5	2.1	1.9	4.7	4.9
	2011	1.7	1.3	0.7	1.2	0.2	0.5	2.0	1.9	4.6	4.8
	2012	1.7	1.2	0.8	1.2	0.2	0.6	1.9	1.8	4.6	4.8
	2013	1.7	1.1	0.8	1.3	0.2	0.6	1.9	1.7	4.6	4.8
	2014	1.7	1.1	0.9	1.4	0.2	0.6	1.8	1.7	4.6	4.8
	2015	1.7	1.1	1.0	1.5	0.2	0.7	1.8	1.6	4.7	4.8
	2016	1.7	1.0	1.0	1.6	0.2	0.7	1.7	1.5	4.7	4.9
	2017	1.7	1.0	1.1	1.7	0.2	0.7	1.7	1.5	4.7	4.9
60	1996	5.8	5.6	-0.3	0.2	-0.6	-0.3	0.8	0.8	5.7	6.3
	1997	5.4	5.1	-0.2	0.3	-0.5	-0.2	0.9	0.9	5.5	6.1
	1998	4.9	4.7	-0.2	0.3	-0.4	-0.2	1.1	1.0	5.4	5.9
	1999	4.5	4.3	-0.1	0.4	-0.4	-0.1	1.3	1.2	5.2	5.7
	2000	4.0	3.9	0.0	0.4	-0.3	-0.1	1.4	1.3	5.1	5.5
	2001	3.7	3.6	0.0	0.4	-0.2	0.0	1.6	1.4	5.0	5.3
	2002	3.3	3.2	0.1	0.5	-0.2	0.0	1.7	1.4	4.9	5.1
	2003	3.0	2.9	0.1	0.5	-0.1	0.1	1.7	1.5	4.7	4.9
	2004	2.8	2.6	0.2	0.5	-0.1	0.1	1.8	1.5	4.6	4.7
	2005	2.5	2.3	0.2	0.6	0.0	0.2	1.8	1.5	4.4	4.6
	2006	2.3	2.0	0.3	0.6	0.0	0.2	1.7	1.5	4.3	4.4
	2007	2.1	1.8	0.3	0.6	0.1	0.3	1.6	1.5	4.1	4.2
	2008	1.9	1.6	0.4	0.7	0.1	0.3	1.5	1.5	3.9	4.0
	2009	1.7	1.4	0.4	0.7	0.2	0.3	1.4	1.4	3.7	3.8
	2010	1.6	1.3	0.5	0.7	0.2	0.4	1.3	1.4	3.6	3.8
	2011	1.6	1.2	0.5	0.8	0.2	0.4	1.3	1.3	3.5	3.7
	2012	1.5	1.1	0.5	0.8	0.2	0.5	1.2	1.3	3.5	3.7
	2013	1.5	1.1	0.6	0.9	0.2	0.5	1.2	1.2	3.5	3.7
	2014	1.5	1.1	0.6	0.9	0.2	0.6	1.1	1.2	3.6	3.7
	2015	1.5	1.0	0.7	1.0	0.2	0.6	1.1	1.1	3.6	3.8
	2016	1.5	1.0	0.8	1.1	0.2	0.6	1.1	1.1	3.6	3.8
	2017	1.6	1.0	0.8	1.2	0.2	0.7	1.1	1.0	3.7	3.8
70	1996	4.7	4.9	-0.3	0.0	-0.5	-0.3	-0.1	-0.1	3.8	4.5
	1997	4.3	4.5	-0.3	0.0	-0.5	-0.3	0.1	0.1	3.7	4.3
	1998	3.9	4.2	-0.2	0.0	-0.4	-0.2	0.3	0.2	3.6	4.2
	1999	3.6	3.9	-0.2	0.0	-0.4	-0.2	0.4	0.3	3.4	4.0
	2000	3.2	3.5	-0.2	0.0	-0.3	-0.2	0.6	0.5	3.3	3.9
	2001	2.9	3.2	-0.1	0.0	-0.2	-0.1	0.7	0.6	3.3	3.7
	2002	2.6	2.9	-0.1	0.1	-0.2	-0.1	0.8	0.7	3.2	3.6
	2003	2.4	2.6	-0.1	0.1	-0.1	0.0	0.9	0.8	3.1	3.4
	2004	2.2	2.3	-0.1	0.1	-0.1	0.0	1.0	0.9	3.0	3.3
	2005	2.0	2.0	0.0	0.2	0.0	0.1	1.0	0.9	3.0	3.2
	2006	1.8	1.8	0.0	0.2	0.0	0.1	1.0	0.9	2.9	3.0

```
2007       1.7  1.6   0.1  0.2   0.1  0.1   0.9  0.9  2.7  2.9
2008       1.5  1.4   0.1  0.2   0.1  0.2   0.9  0.9  2.6  2.7
2009       1.4  1.2   0.1  0.3   0.2  0.2   0.9  0.9  2.5  2.6
2010       1.3  1.1   0.2  0.3   0.2  0.3   0.8  0.9  2.5  2.6
2011       1.2  1.1   0.2  0.3   0.2  0.3   0.8  0.8  2.5  2.5
2012       1.2  1.0   0.2  0.4   0.3  0.3   0.8  0.8  2.5  2.5
2013       1.2  1.0   0.3  0.4   0.3  0.4   0.8  0.8  2.5  2.5
2014       1.2  0.9   0.3  0.5   0.2  0.4   0.8  0.7  2.5  2.6
2015       1.2  0.9   0.3  0.5   0.2  0.4   0.8  0.7  2.6  2.6
2016       1.2  0.9   0.4  0.6   0.2  0.5   0.8  0.7  2.6  2.6
2017       1.2  0.8   0.4  0.7   0.2  0.5   0.8  0.6  2.7  2.7
```

12.6.1 Lifetimne lost by cause of death

```
> str(YLL)
 num [1:2, 1:22, 1:2, 1:100, 1:5] 3.37 6.04 3.69 6.08 4 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
  ..$ A.in: chr [1:100] "0" "1" "2" "3" ...
  ..$ CoD : chr [1:5] "D-CVD" "D-Can" "D-Res" "D-Oth" ...
> cumsum0 <- function(x) cumsum(c(0,x))
> cYLL <- aperm( apply( YLL[,,,,c(1,2,4,3)], 1:4, cumsum0 ), c(2:5,1) )
> str(cYLL)
 num [1:2, 1:22, 1:2, 1:100, 1:5] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
  ..$ A.in: chr [1:100] "0" "1" "2" "3" ...
  ..$     : chr [1:5] "" "D-CVD" "D-Can" "D-Oth" ...
> clr <- adjustcolor( c("red","blue","forestgreen",gray(0.5)), alpha=0.8)
> pl1 <-
+  function( yr=2016, aa=40:100, xl=range(aa), yl=c(0,10.5), sx, tp, inv=(tp=="T1D") )
+  {
+  plot( NA, xlim=if(inv) rev(xl) else xl, ylim=yl,
+            xaxt="n", yaxt="n", xaxs="i", yaxs="i" )
+  yl <<- yl ; xl <<- xl
+  for( i in 1:4 ) filc( aa, cYLL[sx,paste(yr),tp,aa,i+0:1], col=clr[i] )
+  abline( v=4:9*10, h=1:floor(yl[2]), lty="22", lend="butt", col=gray(0.8) )
+  text( mean(xl), yl[2]*0.9, paste(sx,tp,collapse=" ") )
+  box(bty="o",col="black")
+  }
> ltr <- function( ltr ) text( sum(par("usr")[1:2]*c(0.95,0.05)),
+                              sum(par("usr")[3:4]*c(0.05,0.95)), ltr, font=2 )
> ply <-
+  function( yr, yl=c(0,11) )
+  {
+  layout( rbind(1:2,3:4), widths=c(65,65) )
+  par( mar=c(0,0,0,0), oma=c(3,3,1,2), mgp=c(3,1,0)/1.6, las=1 )
+  pl1( yr, sx="M", tp="T1D", aa=35:100, yl=yl ) ; axis( side=2 ) ; ltr("a")
+  pl1( yr, sx="M", tp="T2D", aa=35:100, yl=yl ) ; axis( side=4 ) ; ltr("b")
+  text( sum(xl*c(0.1,0.9)), seq(0.6,0.9,,5)*yl[2],
```

```
+        c( gsub("D-","",dimnames(cYLL)[[5]][-1]), paste(yr) ),
+        col=c(clr,"black"), font=2 )
+  pl1( yr, sx="W", tp="T1D", aa=35:100, yl=yl ) ; axis( side=2 ) ; axis( side=1 ) ; ltr("c")
+  pl1( yr, sx="W", tp="T2D", aa=35:100, yl=yl ) ; axis( side=4 ) ; axis( side=1 ) ; ltr("d")
+  mtext( "Age", side=1, line=2, outer=TRUE )
+  mtext( "Life time lost (years)", side=2, line=2, outer=TRUE, las=0 )
+  }
> ply( 2017 )


> ply( 2017 )


> ply( 2014 )


> ply( 2011, yl=c(0,13) )


> ply( 2008, yl=c(0,13) )


> ply( 2005, yl=c(0,13) )


> ply( 2002, yl=c(0,13) )
```

12.6.2 Total lifetime lost to diabetes

First we plot the overall life time lost to T1 resp. T2 at different ages using cross-sectional rates from date 1996-01-01, 1999-01-01 etc. (every three years). First only the total no. of YLL by sex and type:

```
> clb <- brewer.pal(9,"Blues")[-1]
> clr <- brewer.pal(9,"Reds" )[-1]
> aa <- as.numeric( dimnames(YLL)[["A.in"]] )
> plyll <- function(sx,tp,cl=if(sx=="M") clb else clr,cau="Sum",
+                   xt=(cau!="Sum"),yl=c(0,16))
+     {
+ matplot( NA, ylim=yl, xlim=c(20,90), xlab="", ylab="", yaxs="i", xaxt="n", yaxt="n" )
+ abline( h=-5:20, v=1:9*10, col=gray(0.9), lty="22" )
+ matlines( aa, t(YLL[sx,seq(1,22,3),tp,,cau]),
+           type="l",lty=1, lwd=2, col=cl )
+ text( 80, 13, paste(if(xt) cau,tp,sx), font=2, adj=c(1,1), col=cl[7] )
+ box()
+     }
> par( mfrow=c(2,2), mar=c(0,0,0,0), oma=c(3,3,1,1),
+      mgp=c(3,1,0)/1.6, las=1, bty="o" )
> plyll("M","T1D",clb) ; axis( side=2, col=1)
> plyll("M","T2D",clb)
> plyll("W","T1D",clr) ; axis( side=1) ; axis( side=2)
> plyll("W","T2D",clr) ; axis( side=1)
> mtext( "Age", side=1, line=2, outer=TRUE, cex=0.83 )
> mtext( "Years of life lost to diabetes",
+        side=2, line=2, outer=TRUE, las=0, cex=0.83 )
```

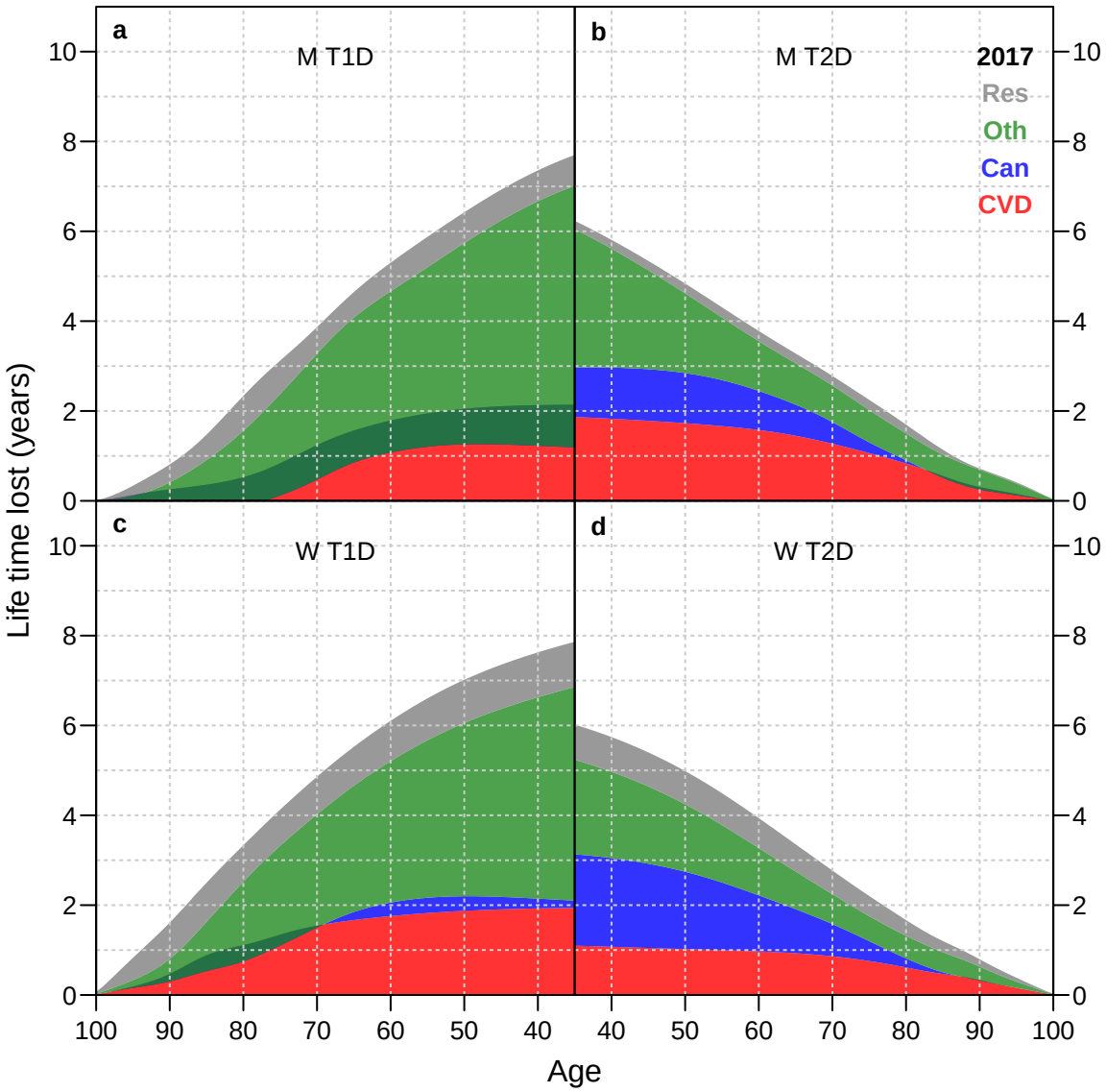


Figure 12.11: *Years of life lost by cause of death in 2017.* `./graph/yll-cod-cau2017`

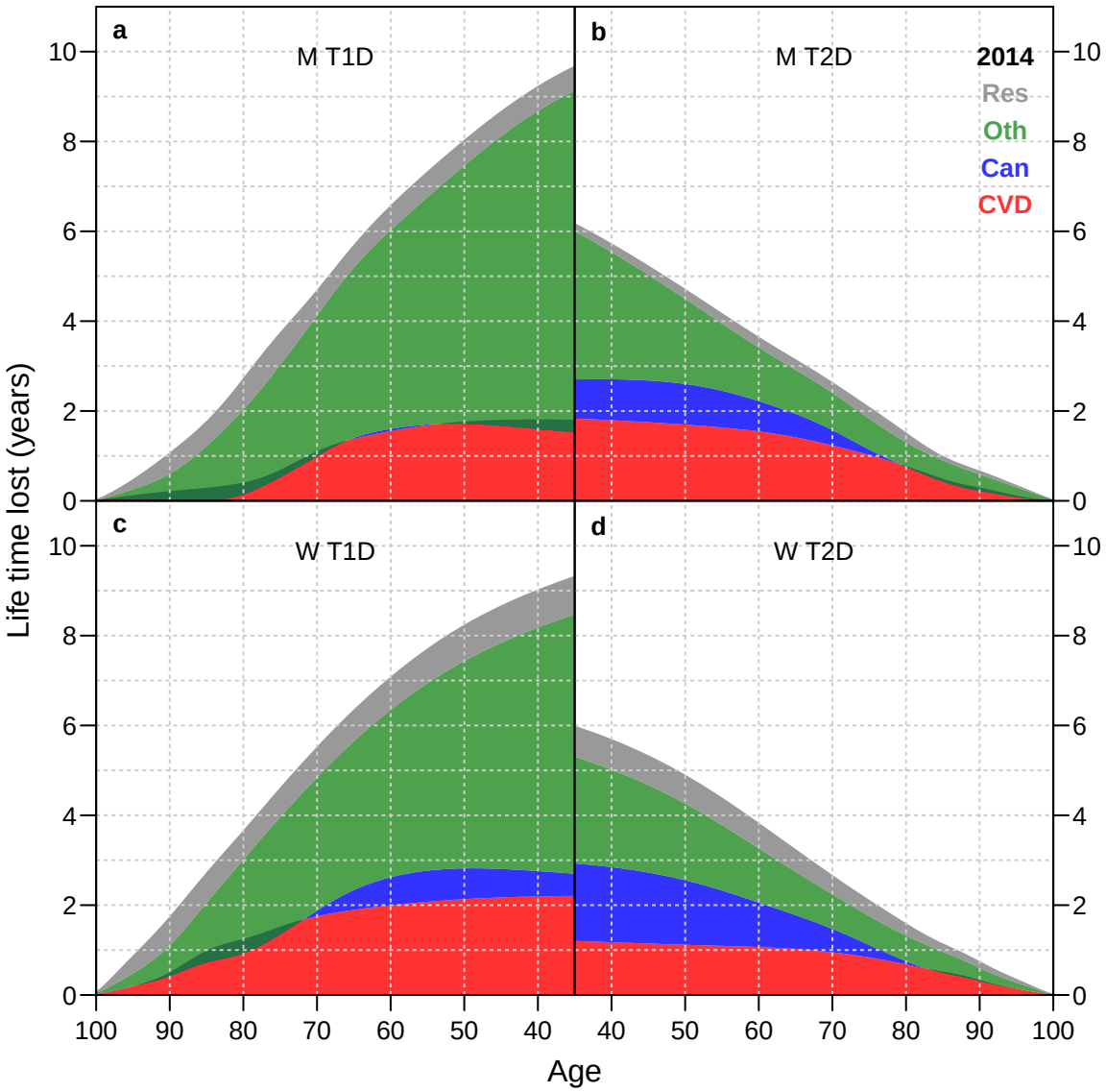


Figure 12.12: *Years of life lost by cause of death in 2014.* `./graph/yll-cod-cau2014`

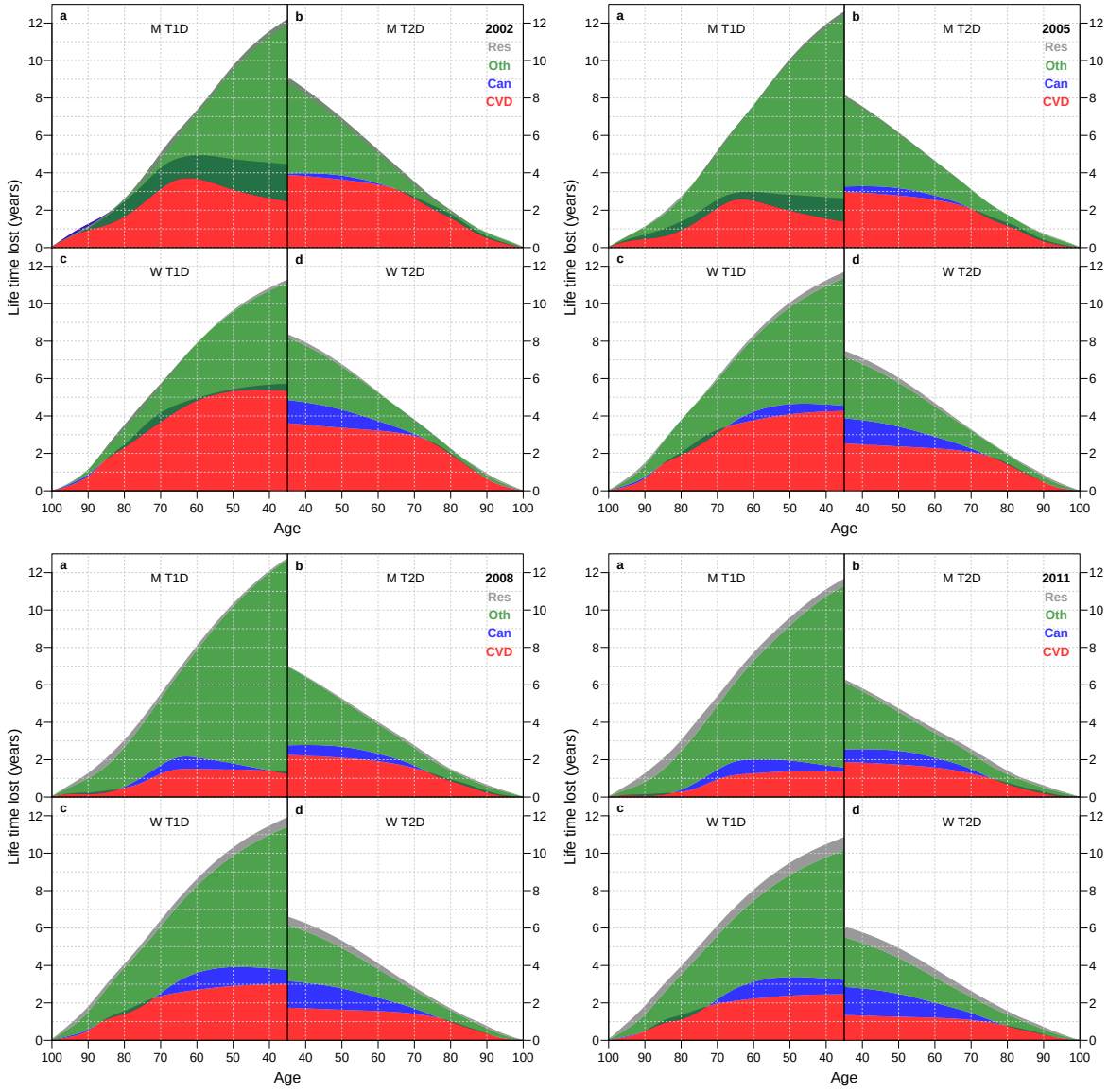


Figure 12.13: *Year of life lost by cause 2002–2011*

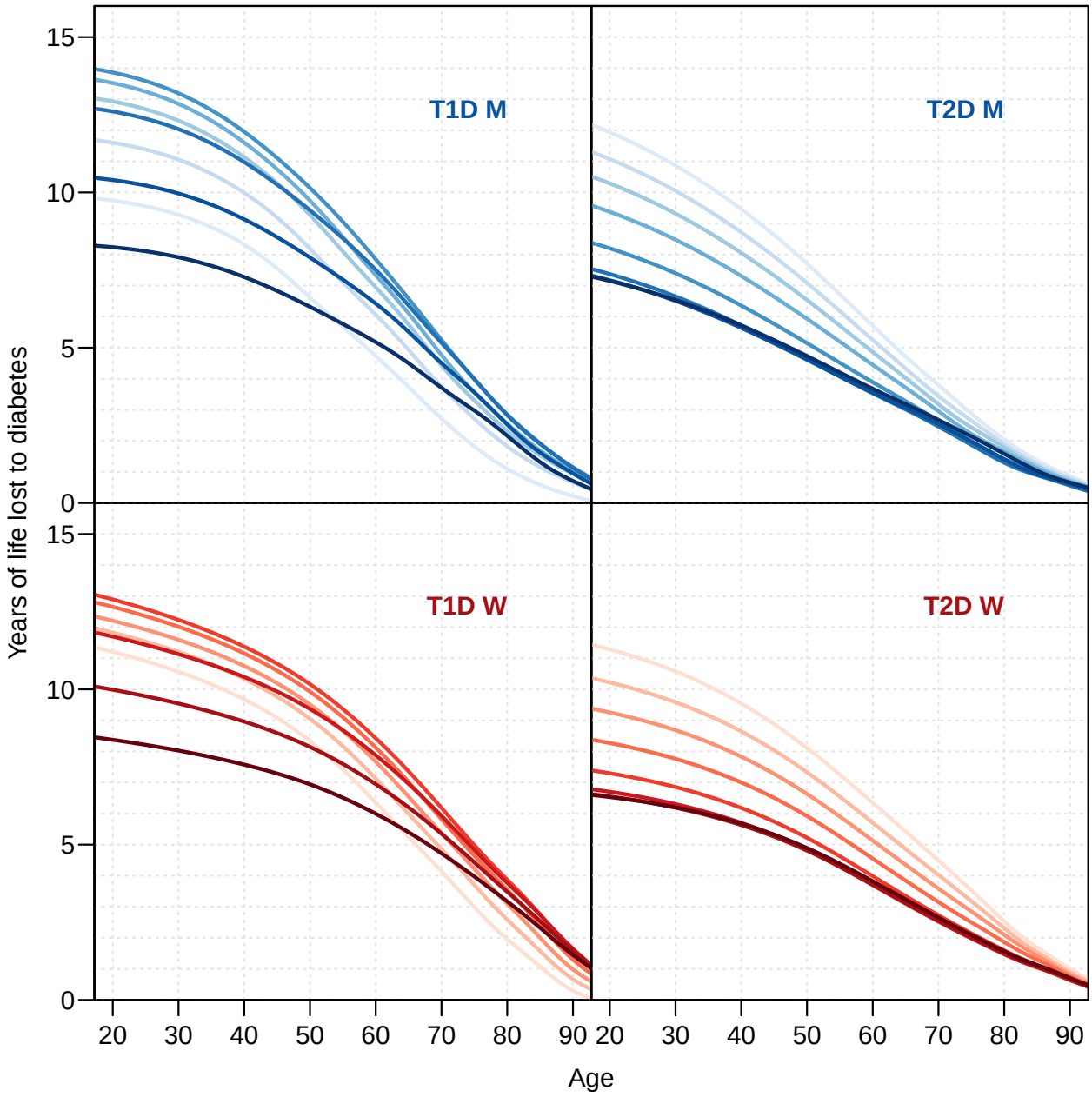


Figure 12.14: *Years of life lost to T1 resp T2 diabetes for men (blue) and women, at 1996,1999,...,2017 (light to dark colour).* `./graph/yll-cod-totYLL`

We can plot the contributions to the years of life lost from different causes as well — this will be a 4 by 4 display:

```
> par( mfcol=c(4,4), mar=c(0,0,0,0), oma=c(3,4,1,1),
+      mgp=c(3,1,0)/1.6, las=1, bty="o" )
> for( tp in c("T1D","T2D") )
+ for( sx in c("M","W") )
+ for( ca in c("D-CVD","D-Can","D-Res","D-Oth") )
+    {
+ plyll(sx,tp,cau=ca,yl=c(-1,14)) ; abline(h=0) ; box()
+ if(sx=="M"&tp=="T1D") axis( side=2, col=1)
+ if(ca=="D-Oth") axis( side=1, col=1)
+    }
> mtext( rep("Age",4), at=(0:3*2+1)/8, side=1, line=2, outer=TRUE, cex=0.66 )
> mtext( "Years of life lost to diabetes from different causes",
+        side=2, line=3, outer=TRUE, las=0, cex=0.66 )
> mtext( c("Other","Respiratory","Cancer","CVD"),
+        side=2, line=2, outer=TRUE, las=0, cex=0.66, at=c(0:3*2+1)/8 )
```

Due to the notorious misclassification of T1 deaths as T2 deaths before 2005, we also make the graphs restricted to the period from 2005 and onward:

```
> aa <- as.numeric( dimnames(YLL)[["A.in"]] )
> plyll <- function(sx,tp,cl=if(sx=="M") clb else clr,cau="Sum",
+                   xt=(cau!="Sum"),yl=c(0,17))
+     {
+ plot( NA, ylim=yl, xlim=c(20,90), xlab="", ylab="", yaxs="i", xaxt="n", yaxt="n" )
+ abline( h=-5:20, v=1:9*10, col=gray(0.8), lty="22" )
+ matlines( aa, t(YLL[sx,yy<-paste(seq(2004,2016,3)),tp,,cau]),
+           type="l", lty=1, lwd=2, col=cl[4:8] )
+ text( 80, 16, paste(if(xt) cau,tp,sx), font=2, adj=c(1,1), col=cl[7] )
+ text( 75.5, 13:9, yy, adj=0, col=cl[5:9] )
+ segments( rep(75,5), 13:9,
+           rep(50,5), YLL[sx,yy,tp,"50",cau], col=cl[4:8] )
+ box()
+     }
> par( mfrow=c(2,2), mar=c(0,0,0,0), oma=c(3,3,1,1),
+      mgp=c(3,1,0)/1.6, las=1, bty="o" )
> plyll("M","T1D",clb) ; axis( side=2, col=1 ) ; box()
> plyll("M","T2D",clb) ; box()
> plyll("W","T1D",clr) ; axis( side=1) ; axis( side=2) ; box()
> plyll("W","T2D",clr) ; axis( side=1) ; box()
> mtext( "Age", side=1, line=2, outer=TRUE, cex=0.83 )
> mtext( "Years of life lost to diabetes",
+        side=2, line=2, outer=TRUE, las=0, cex=0.83 )


> par( mfcol=c(4,4), mar=c(0.1,0.1,0.1,0.1)*0, oma=c(3,4,1,1),
+      mgp=c(3,1,0)/1.6, las=1, bty="o" )
> for( tp in c("T1D","T2D") )
+ for( sx in c("M","W") )
+ for( ca in c("D-CVD","D-Can","D-Res","D-Oth") )
+    {
+ plyll(sx,tp,cau=ca,yl=c(-4,18),xt=FALSE)
+ abline(h=0)
+ if(sx=="M"&tp=="T1D") axis( side=2, col=1)
+ if(ca=="D-Oth") axis( side=1, col=1)
+    }
```

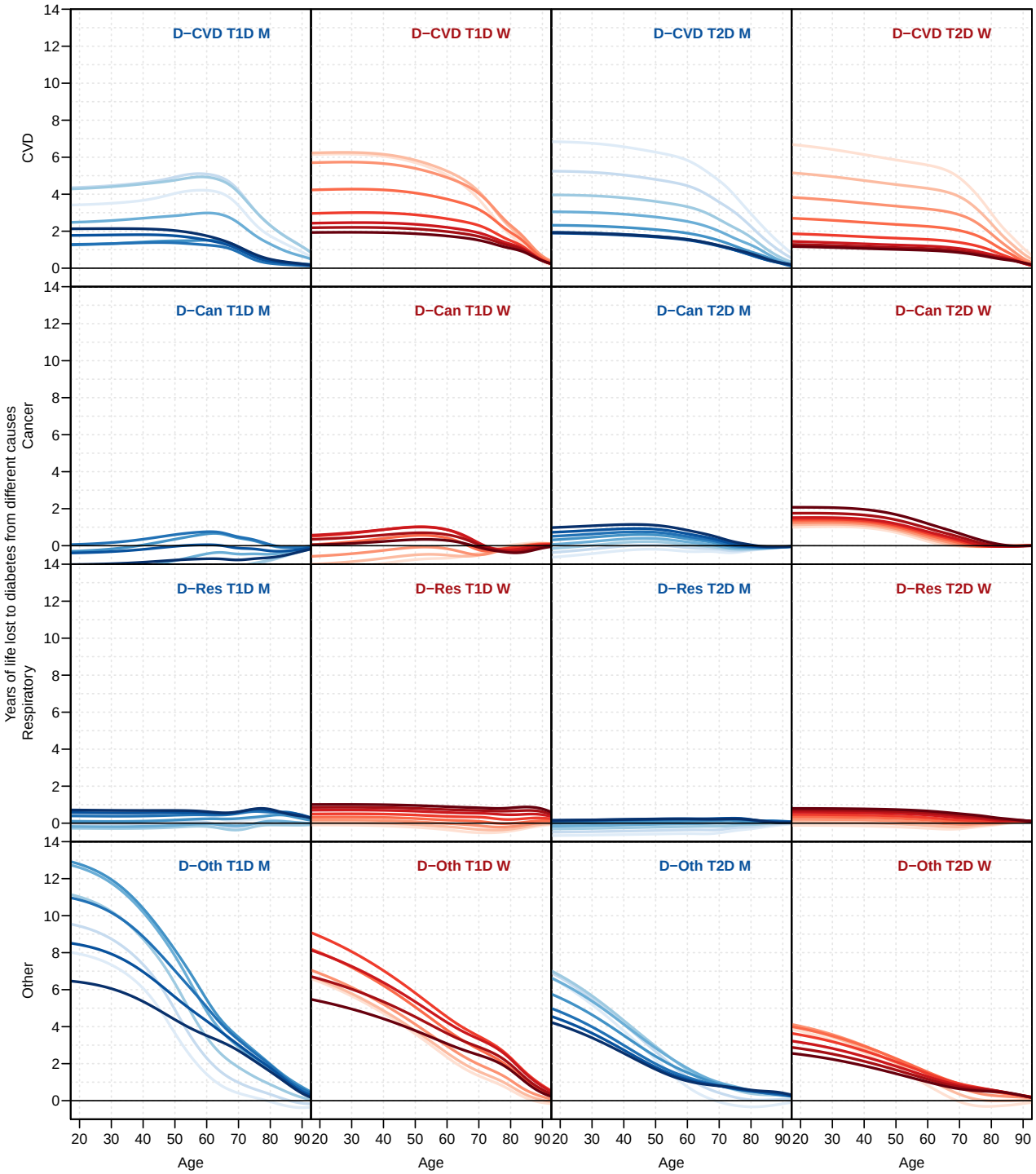


Figure 12.15: *Years of life lost to T1 resp T2 diabetes subdivided by cause for men (blue) and women, each year 1996, 1999,. . . ,2017 (light to dark colour).* `./graph/yll-cod-allYLL`

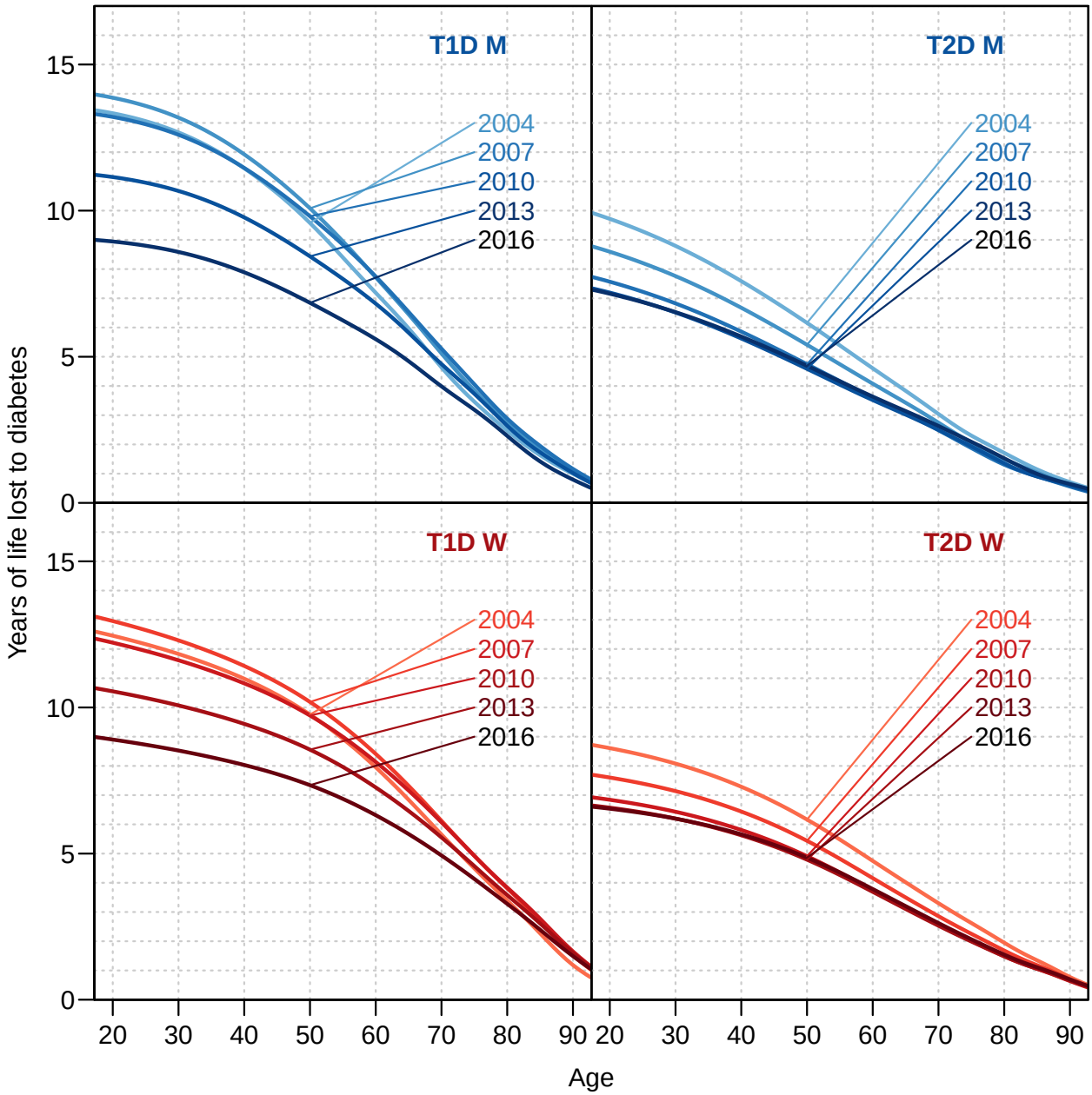


Figure 12.16: *Years of life lost to T1 resp T2 diabetes for men (blue) and women, at 2005,2008,...,2017 (light to dark colour).* `./graph/yll-cod-totYLLlate`

```
> mtext( "Age", side=1, line=2, outer=TRUE, cex=0.66 )
> mtext( "Years of life lost to diabetes from different causes",
+        side=2, line=3, outer=TRUE, las=0, cex=0.66 )
> mtext( c("Other","Respiratory","Cancer","CVD"),
+        side=2, line=2, outer=TRUE, las=0, cex=0.66, at=c(0:3*2+1)/8 )
```

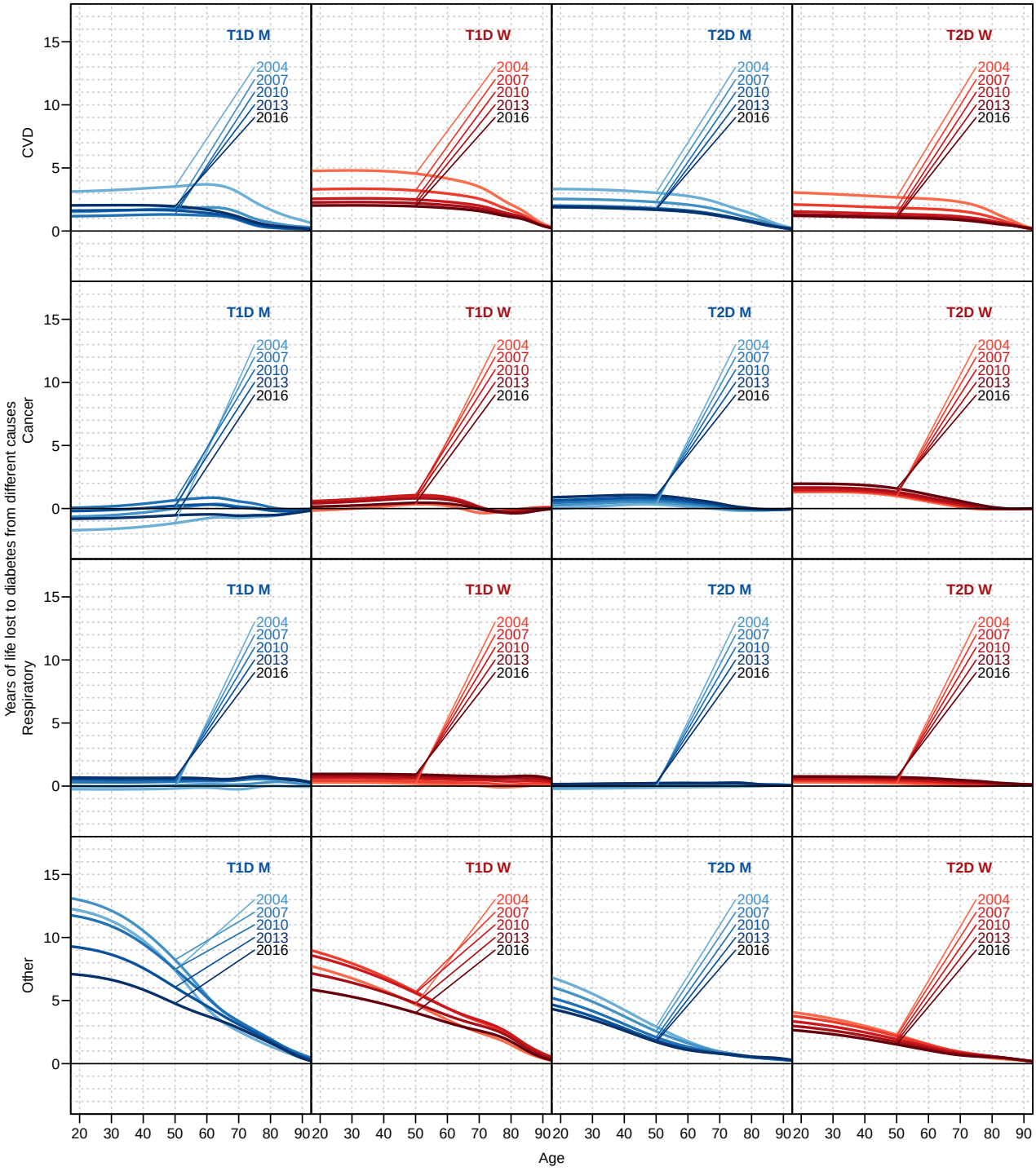


Figure 12.17: *Years of life lost to T1 resp T2 diabetes subdivided by cause for men (blue) and women, at 2005,2008,...,2017 (light to dark colour).* `./graph/yll-cod-allYLLlate`

We note that the the years of life lost need not be positive, neither in total or for the specific causes of death. We see that we have negative lifetime lost to cancer for T1D patients, indicating that these patients are more likely to die from other causes early in life and thus escaping cancer death.

12.7 Population burden of diabetes

There are several ways of quantifying the population burden of diabetes in terms of lifetime lost; either using the prevelant cases at at given date or using the incident cases over a given period, and either using some truncated lifespan or the entire (anticipated) lifespan.

In order to compute what we could call the *annual* burden of diabetes in the Danish population we will use the 1-year restricted years of life lost at all ages. In practise this will be in 1-year intervals of age, using the midpoint of each age-group as starting age and only integratin over a period of 1 year. These numbers will then be multiplied by the number of prevalent cases at 1 January each year, to yield the total population lifetime lost each year.

This will be the number of years additionally lived during a year if all diabetes patients reverted to non-DM mortality rates. As we shall see this is going to be quite a small quantity, because we are comparing two survival curves that are identical at onset — we will only ever be comparing curves that are very close.

Another more realistic approach would be to compute the years of life lost over the rest of all patients' lives, using prevalent cases of T1D/T2D at a given date. Showing this over time will illustrate to what extent the future burden is increasing.

A similar measure could be computed using only the annual number of *new* cases — the incident cases. Showing this over time would give an impression of the future burden of diabetes added each year.

Both of these two latter measures would require using the `YLL` array and then just multiplying it with the number or incident respectively prevalent cases.

12.7.1 Diabetes population

Prevalent cases

The prevalent cases we already have from the prevalence analysis and chapter 4, and they are easily put in the tabular form that will be used for the one-year lifelost as well.

```
> load( "../nydata/prevN.Rda" )
> str( prN )
'data.frame':        22000 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","W": 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 2 3 2 0 3 6 1 3 ...
 $ T2 : num  0 0 0 0 0 0 0 0 0 0 ...
 $ nD : num  3822 3939 3930 3872 3685 ...
 $ N  : num  3822 3939 3932 3875 3687 ...
 $ DM : num  0 0 2 3 2 0 3 6 1 3 ...
> pr <- aperm( xtabs( cbind(T1,T2) ~ sex + A + P, data=prN ),
+              c(1,3,4,2) )
> str( pr )
```

```
 'xtabs' num [1:2, 1:22, 1:2, 1:100] 0 0 0 0 0 0 0 0 1 0 ...
 - attr(*, "dimnames")=List of 4
  ..$ sex: chr [1:2] "M" "W"
  ..$ P  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$    : chr [1:2] "T1" "T2"
  ..$ A  : chr [1:100] "0" "1" "2" "3" ...
> fCtable( addmargins( apply(pr,1:3,sum), c(1,3) ), col.vars=c(3,1), w=7 )
             T1                      T2                     Sum
     sex      M       W     Sum       M       W     Sum       M       W     Sum
P
1996      12,328   9,549  21,877  30,266  31,303  61,569  42,594  40,852  83,446
1997      12,677   9,776  22,453  33,788  34,100  67,888  46,465  43,876  90,341
1998      12,958   9,986  22,944  36,950  36,423  73,373  49,908  46,409  96,317
1999      13,222  10,113  23,335  40,709  39,148  79,857  53,931  49,261 103,192
2000      13,386  10,235  23,621  44,394  42,118  86,512  57,780  52,353 110,133
2001      13,560  10,295  23,855  47,954  44,890  92,844  61,514  55,185 116,699
2002      13,729  10,371  24,100  51,621  47,461  99,082  65,350  57,832 123,182
2003      13,845  10,450  24,295  56,322  51,802 108,124  70,167  62,252 132,419
2004      13,948  10,477  24,425  61,903  56,391 118,294  75,851  66,868 142,719
2005      14,011  10,566  24,577  67,635  61,095 128,730  81,646  71,661 153,307
2006      14,072  10,643  24,715  72,155  64,317 136,472  86,227  74,960 161,187
2007      14,208  10,715  24,923  76,551  66,924 143,475  90,759  77,639 168,398
2008      14,339  10,801  25,140  81,384  70,287 151,671  95,723  81,088 176,811
2009      14,485  10,901  25,386  87,369  74,558 161,927 101,854  85,459 187,313
2010      14,648  10,979  25,627  93,772  78,757 172,529 108,420  89,736 198,156
2011      14,745  11,076  25,821 101,215  83,728 184,943 115,960  94,804 210,764
2012      14,860  11,176  26,036 112,076  93,091 205,167 126,936 104,267 231,203
2013      14,988  11,289  26,277 119,920  99,319 219,239 134,908 110,608 245,516
2014      15,116  11,458  26,574 125,068 103,279 228,347 140,184 114,737 254,921
2015      15,304  11,614  26,918 129,576 106,519 236,095 144,880 118,133 263,013
2016      15,512  11,826  27,338 134,157 109,769 243,926 149,669 121,595 271,264
2017      15,684  11,930  27,614 139,192 113,232 252,424 154,876 125,162 280,038
```

Incident cases

The incident cases are available in the `rt` data frame

```
> str( rt )
'data.frame':        354350 obs. of  18 variables:
 $ sex  : Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ state: Factor w/ 3 levels "noDM","T1","T2": 2 2 2 2 2 2 2 2 2 2 ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ P    : num  1997 1999 1999 2000 2000 ...
 $ C    : num  1996 1998 1999 1999 1999 ...
 $ dur  : num  0.1 0.1 0.1 0.1 0.35 0.1 0.1 0.1 0.35 0.1 ...
 $ Dcvd : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dcan : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dres : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Doth : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y    : num  1.43e-04 2.26e-05 7.67e-05 3.79e-04 1.76e-04 ...
 $ T1   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Ax   : num  0.667 0.667 0.333 0.667 0.667 ...
 $ Px   : num  1997 1999 2000 2000 2000 ...
 $ Cx   : num  1997 1999 1999 2000 2000 ...
 $ gP   : Factor w/ 7 levels "1996-1998","1999-2001",..: 1 2 2 2 2 2 3 3 3 3 ...
```

```
> ic <- aperm( xtabs( cbind(T1,T2) ~ sex + A + P, data=rt ),
+              c(1,3,4,2) )
> str( ic )
 'xtabs' num [1:2, 1:21, 1:2, 1:100] 0 3 1 2 0 0 2 1 2 0 ...
 - attr(*, "dimnames")=List of 4
  ..$ sex: chr [1:2] "M" "W"
  ..$ P  : chr [1:21] "1996" "1997" "1998" "1999" ...
  ..$    : chr [1:2] "T1" "T2"
  ..$ A  : chr [1:100] "0" "1" "2" "3" ...
> fCtable( addmargins( apply(ic,1:3,sum), c(1,3) ), col.vars=c(3,1), w=6 )
             T1                   T2                  Sum
     sex      M      W    Sum      M      W    Sum      M      W    Sum
P
1996         678    516  1,194  6,115  5,290 11,405  6,793  5,806 12,599
1997         684    489  1,173  5,839  4,918 10,757  6,523  5,407 11,930
1998         657    454  1,111  6,529  5,295 11,824  7,186  5,749 12,935
1999         592    413  1,005  6,739  5,707 12,446  7,331  6,120 13,451
2000         596    392    988  6,593  5,604 12,197  7,189  5,996 13,185
2001         586    415  1,001  6,795  5,449 12,244  7,381  5,864 13,245
2002         602    386    988  8,022  7,334 15,356  8,624  7,720 16,344
2003         545    386    931  9,146  7,673 16,819  9,691  8,059 17,750
2004         509    388    897  9,259  7,751 17,010  9,768  8,139 17,907
2005         517    379    896  8,174  6,510 14,684  8,691  6,889 15,580
2006         554    382    936  8,172  5,940 14,112  8,726  6,322 15,048
2007         564    384    948  8,738  6,792 15,530  9,302  7,176 16,478
2008         546    367    913  9,846  7,554 17,400 10,392  7,921 18,313
2009         568    357    925 10,762  7,720 18,482 11,330  8,077 19,407
2010         529    367    896 11,867  8,704 20,571 12,396  9,071 21,467
2011         496    358    854 15,593 13,150 28,743 16,089 13,508 29,597
2012         486    315    801 12,782 10,017 22,799 13,268 10,332 23,600
2013         471    351    822 10,215  7,971 18,186 10,686  8,322 19,008
2014         465    341    806  9,883  7,358 17,241 10,348  7,699 18,047
2015         476    375    851  9,987  7,638 17,625 10,463  8,013 18,476
2016         460    316    776 10,666  7,855 18,521 11,126  8,171 19,297
```

12.7.2 Total fututre lifetime lost

This is merely a question of multiplying the `YLL` array of future lifetime lost with the right numbers of diabetes persons:

```
> str( YLL )
 num [1:2, 1:22, 1:2, 1:100, 1:5] 3.37 6.04 3.69 6.08 4 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
  ..$ A.in: chr [1:100] "0" "1" "2" "3" ...
  ..$ CoD : chr [1:5] "D-CVD" "D-Can" "D-Res" "D-Oth" ...
```

Prevalent cases

Here we multiply with the prevalent number of cases which means that we are incuding years lost to diabetes from the same persons in succesive years.

```
> str( pr )
 'xtabs' num [1:2, 1:22, 1:2, 1:100] 0 0 0 0 0 0 0 0 1 0 ...
 - attr(*, "dimnames")=List of 4
  ..$ sex: chr [1:2] "M" "W"
  ..$ P  : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$    : chr [1:2] "T1" "T2"
  ..$ A  : chr [1:100] "0" "1" "2" "3" ...
> YLLpr <- YLL * 0
> for( i in 1:5 ) YLLpr[,,,,i] <- YLL[,,,,i] * pr
> YLLpr <- apply( YLLpr, c(1:3,5), sum )
> str( YLLpr )
 num [1:2, 1:22, 1:2, 1:5] 44609 48608 49856 50560 54989 ...
 - attr(*, "dimnames")=List of 4
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
  ..$ CoD : chr [1:5] "D-CVD" "D-Can" "D-Res" "D-Oth" ...
> fCtable( addmargins( YLLpr[,seq(1,22,3),,5], c(1,3) ),
+          row.vars=3:2, d=1, w=11 )
          sex           M           W         Sum
S.in date
T1D  1996        85,077.0    74,696.7   159,773.7
     1999       110,933.0    86,647.6   197,580.6
     2002       129,833.5    94,757.8   224,591.3
     2005       138,836.4   102,209.0   241,045.5
     2008       148,128.6   108,545.6   256,674.3
     2011       141,362.5   103,236.6   244,599.1
     2014       121,203.0    92,669.7   213,872.6
     2017        99,899.6    81,930.7   181,830.3
T2D  1996       149,370.2   150,603.1   299,973.3
     1999       188,227.4   173,540.9   361,768.3
     2002       222,776.4   192,579.1   415,355.5
     2005       266,763.1   223,320.0   490,083.1
     2008       276,754.4   225,409.6   502,164.0
     2011       310,471.1   248,444.8   558,915.9
     2014       377,869.7   303,678.1   681,547.8
     2017       428,532.2   337,258.6   765,790.8
Sum  1996       234,447.2   225,299.8   459,747.0
     1999       299,160.4   260,188.5   559,348.9
     2002       352,609.9   287,336.9   639,946.9
     2005       405,599.5   325,529.0   731,128.5
     2008       424,883.0   333,955.2   758,838.2
     2011       451,833.6   351,681.4   803,515.0
     2014       499,072.7   396,347.8   895,420.5
     2017       528,431.8   419,189.3   947,621.1
```

We then plot the annually updated future burden of diabetes:

```
> par( mar=c(3,4,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> fYLL <- function(){
+ plot( NA,
+       xlim=c(1996,2017),
+       ylim=c(0,450), yaxs="i",
+       xlab="Date of prevalence",
+       ylab="" )
+ mtext( "Future life years lost (1000s) to diabetes",
```

```
+         side=2, line=3, las=0, cex=1-0.17*(all(par('mfcol')==c(2,2))) )
+ abline( h=1:9*50, col=gray(0.9) )
+ axis( side=1, at=1996:2017, tcl=-0.3, labels=NA )
+ axis( side=2, at=0:9*50, tcl=-0.3, labels=NA )
+ matlines( as.numeric(dimnames(YLLpr)[[2]]),
+           cbind( YLLpr["M",,"T1D","Sum"],
+                  YLLpr["M",,"T2D","Sum"],
+                  YLLpr["W",,"T1D","Sum"],
+                  YLLpr["W",,"T2D","Sum"] ) / 1000,
+          col=rep(c("blue","red"),each=2), lty=c("21","solid"),
+          lend="butt", lwd=4, type="l" )
+ }
> fYLL()
```

We can also compute the *average* number of future years lost among the prevalent diabetes patients; that is the *per person* future burden of diabetes for those alive at a given date:

```
> fCtable( YLLpr.p <- addmargins( YLLpr[,,,5], 1 ) /
+                     addmargins( apply( pr, 1:3, sum ), 1 ),
+          w=5, d=1, col.vars=c(3,1) )
     S.in   T1D               T2D
     sex      M     W   Sum     M     W   Sum
date
1996         6.9   7.8   7.3   4.9   4.8   4.9
1997         7.4   8.1   7.7   4.8   4.7   4.8
1998         7.9   8.3   8.1   4.7   4.6   4.6
1999         8.4   8.6   8.5   4.6   4.4   4.5
2000         8.8   8.8   8.8   4.5   4.3   4.4
2001         9.2   9.0   9.1   4.4   4.2   4.3
2002         9.5   9.1   9.3   4.3   4.1   4.2
2003         9.6   9.3   9.5   4.2   3.9   4.1
2004         9.8   9.5   9.6   4.1   3.8   4.0
2005         9.9   9.7   9.8   3.9   3.7   3.8
2006        10.1   9.9  10.0   3.8   3.5   3.6
2007        10.3  10.0  10.2   3.6   3.3   3.5
2008        10.3  10.0  10.2   3.4   3.2   3.3
2009        10.2   9.9  10.1   3.2   3.1   3.2
2010        10.0   9.7   9.8   3.1   3.0   3.1
2011         9.6   9.3   9.5   3.1   3.0   3.0
2012         9.1   8.9   9.0   3.0   2.9   3.0
2013         8.6   8.5   8.5   3.0   2.9   3.0
2014         8.0   8.1   8.0   3.0   2.9   3.0
2015         7.5   7.7   7.6   3.0   2.9   3.0
2016         6.9   7.3   7.1   3.1   3.0   3.0
2017         6.4   6.9   6.6   3.1   3.0   3.0
> YLLpr.x <- YLLpr.p[,,rep(1:2,c(4,1))]
> dimnames( YLLpr.x )[[3]] <- c("T1D","T2D","DM","T1D/pp","T2D/pp")
> YLLpr.x[,,1:3] <- addmargins( YLLpr[,,,5], c(1,3) )/1000
> fCtable( YLLpr.x[,seq(1,22,3),], w=10, d=1, row.vars=1:2 )
         S.in        T1D        T2D         DM     T1D/pp     T2D/pp
sex date
M   1996           85.1      149.4      234.4        6.9        4.9
    1999          110.9      188.2      299.2        8.4        4.6
    2002          129.8      222.8      352.6        9.5        4.3
    2005          138.8      266.8      405.6        9.9        3.9
    2008          148.1      276.8      424.9       10.3        3.4
    2011          141.4      310.5      451.8        9.6        3.1
```

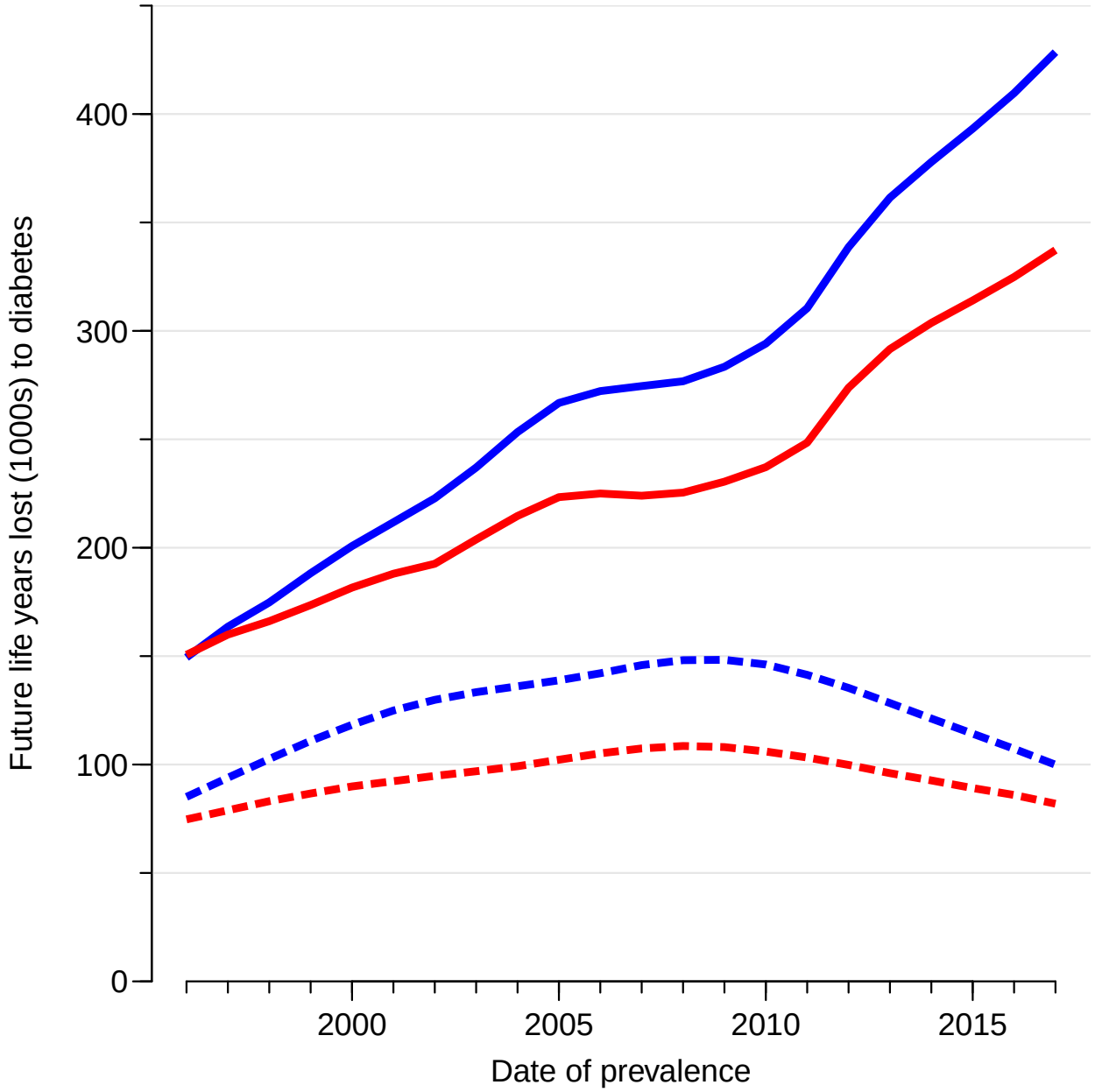


Figure 12.18: *Danish population total of future years of life lost to diabetes in* prevalent *patients alive 1 January 1996–2017 — the total future years of life lost among the diabetes patients alive at each date. Blue lines are men, red women, dotted lines are T1D and full lines T2D. Note that for each year some of the years lost are counted in subsequent years too.* `./graph/yll-cod-YLLpr`

	2014	121.2	377.9	499.1	8.0	3.0
	2017	99.9	428.5	528.4	6.4	3.1
W	1996	74.7	150.6	225.3	7.8	4.8
	1999	86.6	173.5	260.2	8.6	4.4
	2002	94.8	192.6	287.3	9.1	4.1
	2005	102.2	223.3	325.5	9.7	3.7
	2008	108.5	225.4	334.0	10.0	3.2
	2011	103.2	248.4	351.7	9.3	3.0
	2014	92.7	303.7	396.3	8.1	2.9

`yll-cod.tex`

```
     2017          81.9      337.3      419.2       6.9       3.0
Sum  1996         159.8      300.0      459.7       7.3       4.9
     1999         197.6      361.8      559.3       8.5       4.5
     2002         224.6      415.4      639.9       9.3       4.2
     2005         241.0      490.1      731.1       9.8       3.8
     2008         256.7      502.2      758.8      10.2       3.3
     2011         244.6      558.9      803.5       9.5       3.0
     2014         213.9      681.5      895.4       8.0       3.0
     2017         181.8      765.8      947.6       6.6       3.0
> fCtable( YLLpr.x, w=10, d=1, row.vars=1:2 )
          S.in        T1D        T2D         DM     T1D/pp     T2D/pp
sex  date
M    1996          85.1      149.4      234.4       6.9       4.9
     1997          93.9      163.6      257.5       7.4       4.8
     1998         102.6      174.8      277.3       7.9       4.7
     1999         110.9      188.2      299.2       8.4       4.6
     2000         118.3      200.8      319.1       8.8       4.5
     2001         124.9      211.7      336.6       9.2       4.4
     2002         129.8      222.8      352.6       9.5       4.3
     2003         133.4      237.0      370.3       9.6       4.2
     2004         136.1      253.3      389.4       9.8       4.1
     2005         138.8      266.8      405.6       9.9       3.9
     2006         142.1      272.2      414.3      10.1       3.8
     2007         145.9      274.5      420.4      10.3       3.6
     2008         148.1      276.8      424.9      10.3       3.4
     2009         148.3      283.4      431.7      10.2       3.2
     2010         146.1      294.1      440.3      10.0       3.1
     2011         141.4      310.5      451.8       9.6       3.1
     2012         135.3      338.6      473.9       9.1       3.0
     2013         128.4      361.4      489.8       8.6       3.0
     2014         121.2      377.9      499.1       8.0       3.0
     2015         114.2      393.3      507.5       7.5       3.0
     2016         107.2      409.6      516.9       6.9       3.1
     2017          99.9      428.5      528.4       6.4       3.1
W    1996          74.7      150.6      225.3       7.8       4.8
     1997          78.9      159.9      238.8       8.1       4.7
     1998          83.1      166.1      249.2       8.3       4.6
     1999          86.6      173.5      260.2       8.6       4.4
     2000          89.9      181.6      271.5       8.8       4.3
     2001          92.3      188.0      280.3       9.0       4.2
     2002          94.8      192.6      287.3       9.1       4.1
     2003          96.9      203.8      300.7       9.3       3.9
     2004          99.2      214.6      313.8       9.5       3.8
     2005         102.2      223.3      325.5       9.7       3.7
     2006         105.2      225.0      330.1       9.9       3.5
     2007         107.4      224.0      331.4      10.0       3.3
     2008         108.5      225.4      334.0      10.0       3.2
     2009         108.0      230.5      338.5       9.9       3.1
     2010         106.0      237.1      343.1       9.7       3.0
     2011         103.2      248.4      351.7       9.3       3.0
     2012          99.8      273.8      373.6       8.9       2.9
     2013          96.0      291.6      387.6       8.5       2.9
     2014          92.7      303.7      396.3       8.1       2.9
     2015          89.1      314.0      403.2       7.7       2.9
     2016          86.0      324.8      410.8       7.3       3.0
     2017          81.9      337.3      419.2       6.9       3.0
Sum  1996         159.8      300.0      459.7       7.3       4.9
```

```
1997          172.8      323.5      496.3        7.7        4.8
1998          185.7      340.9      526.5        8.1        4.6
1999          197.6      361.8      559.3        8.5        4.5
2000          208.2      382.4      590.6        8.8        4.4
2001          217.2      399.7      616.9        9.1        4.3
2002          224.6      415.4      639.9        9.3        4.2
2003          230.3      440.8      671.0        9.5        4.1
2004          235.3      468.0      703.3        9.6        4.0
2005          241.0      490.1      731.1        9.8        3.8
2006          247.2      497.2      744.5       10.0        3.6
2007          253.3      498.5      751.8       10.2        3.5
2008          256.7      502.2      758.8       10.2        3.3
2009          256.3      513.9      770.2       10.1        3.2
2010          252.1      531.3      783.4        9.8        3.1
2011          244.6      558.9      803.5        9.5        3.0
2012          235.2      612.4      847.5        9.0        3.0
2013          224.4      653.1      877.4        8.5        3.0
2014          213.9      681.5      895.4        8.0        3.0
2015          203.4      707.3      910.6        7.6        3.0
2016          193.2      734.5      927.7        7.1        3.0
2017          181.8      765.8      947.6        6.6        3.0
```

. . . and make a plot of these average numbers; these are shown in figure 12.19.

```
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> fYLLpp <- function(){
+ plot( NA,
+       xlim=c(1996,2017),
+       ylim=c(0,13), yaxs="i",
+       xlab="Date of prevalence",
+       ylab="Average years lost to diabetes per patient")
+ abline( h=1:15, col=gray(0.9) )
+ axis( side=1, at=1996:2017, tcl=-0.2, labels=NA )
+ # axis( side=2, at=0:6*2, tcl=-0.3, labels=NA )
+ axis( side=2, at=0:13 , tcl=-0.3, labels=NA )
+ matlines( as.numeric(dimnames(YLLpr.p)[[2]]),
+           cbind( YLLpr.p["M",,"T1D"],
+                  YLLpr.p["M",,"T2D"],
+                  YLLpr.p["W",,"T1D"],
+                  YLLpr.p["W",,"T2D"] ),
+          col=rep(c("blue","red"),each=2), lty=c("21","solid"),
+          lend="butt", lwd=4, type="l" )
+ }
> fYLLpp()
```

Incident cases

Here we multiply the years of life lost classified by sex type and age with the incident number of cases each year, and sum it over ages. This means that we are incuding years lost to diabetes in the future from the newly entered persons in each year. Note that we are now referring to persons collected over a *period* of a year, so we use the average of the YLL from the first and last day of the year:

```
> str( YLL )
```

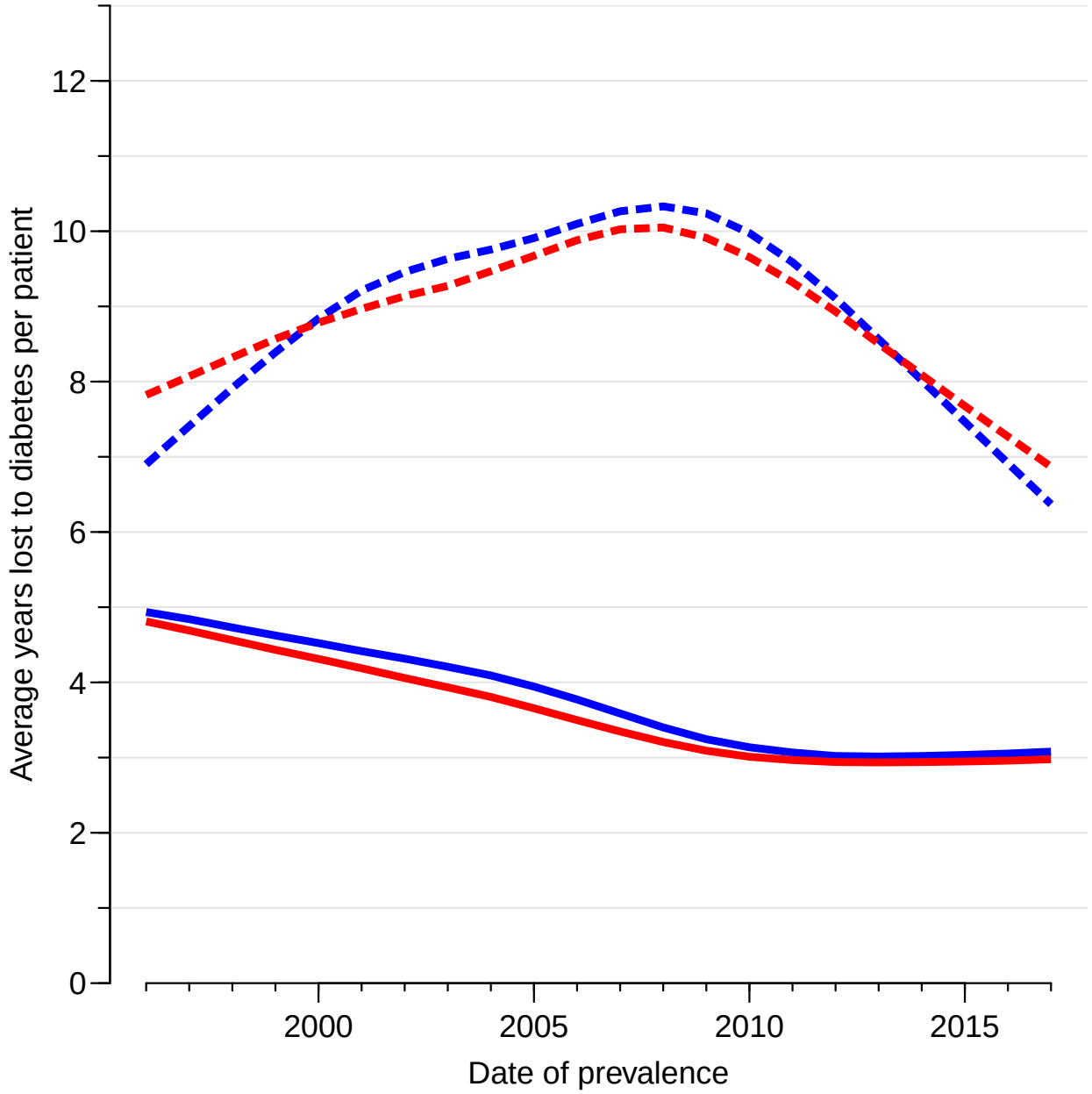


Figure 12.19: *Future years of life lost to diabetes* per patient *alive each year 1996–2017. Blue lines are men, red women, dotted lines are T1D and full lines T2D. Note again, that since the figures refer to* prevalent *cases, numbers from one year are part of subsequent years as well.*
`./graph/yll-cod-YLLpr-pp`

```
 num [1:2, 1:22, 1:2, 1:100, 1:5] 3.37 6.04 3.69 6.08 4 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
  ..$ A.in: chr [1:100] "0" "1" "2" "3" ...
  ..$ CoD : chr [1:5] "D-CVD" "D-Can" "D-Res" "D-Oth" ...
> YLLic <- YLL[,-dim(YLL)[2],,,] * 0
> str( YLLic )
```

```
 num [1:2, 1:21, 1:2, 1:100, 1:5] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:21] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
  ..$ A.in: chr [1:100] "0" "1" "2" "3" ...
  ..$ CoD : chr [1:5] "D-CVD" "D-Can" "D-Res" "D-Oth" ...
```

```
> for( i in 1:5 ) YLLic[,,,,i] <- (YLL[,-1          ,,,i]+
+                                  YLL[,-dim(YLL)[2],,,i])/2 * ic
> YLLic <- apply( YLLic, c(1,2,3,5), sum )
> str( YLLic )
```

```
 num [1:2, 1:21, 1:2, 1:5] 2536 2692 2748 2675 2827 ...
 - attr(*, "dimnames")=List of 4
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:21] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
  ..$ CoD : chr [1:5] "D-CVD" "D-Can" "D-Res" "D-Oth" ...
```

```
> fCtable( addmargins( YLLic[,seq(3,21,3),,5], c(1,3) ),
+          row.vars=1:2, d=1, w=10 )
```

	S.in	T1D	T2D	Sum
sex	date			
M	1998	5,757.8	33,376.8	39,134.6
	2001	6,096.4	32,815.8	38,912.3
	2004	5,749.2	40,636.7	46,385.9
	2007	6,652.9	34,104.2	40,757.1
	2010	6,079.2	41,343.6	47,422.8
	2013	4,703.5	34,857.4	39,560.9
	2016	3,748.0	38,619.8	42,367.8
W	1998	4,335.4	26,082.7	30,418.1
	2001	4,334.4	24,678.6	29,012.9
	2004	4,180.1	32,055.8	36,236.0
	2007	4,419.4	25,293.3	29,712.8
	2010	4,090.7	29,556.3	33,647.0
	2013	3,421.8	26,917.2	30,339.0
	2016	2,652.9	28,177.2	30,830.1
Sum	1998	10,093.2	59,459.5	69,552.7
	2001	10,430.8	57,494.4	67,925.2
	2004	9,929.3	72,692.5	82,621.9
	2007	11,072.3	59,397.5	70,469.8
	2010	10,169.9	70,899.9	81,069.8
	2013	8,125.3	61,774.6	69,899.9
	2016	6,400.9	66,797.0	73,197.9

```
> fCtable( addmargins( YLLic[,,,5], c(1,3) ),
+          row.vars=1:2, d=1, w=10 )
```

	S.in	T1D	T2D	Sum
sex	date			
M	1996	5,270.1	32,491.9	37,762.0
	1997	5,638.6	30,367.6	36,006.2
	1998	5,757.8	33,376.8	39,134.6
	1999	5,720.0	33,663.7	39,383.7
	2000	5,997.0	32,285.4	38,282.4
	2001	6,096.4	32,815.8	38,912.3
	2002	6,464.9	37,174.5	43,639.4
	2003	5,858.6	41,560.5	47,419.0
	2004	5,749.2	40,636.7	46,385.9

	2005	6,079.0	35,035.6	41,114.7
	2006	6,510.9	33,367.7	39,878.6
	2007	6,652.9	34,104.2	40,757.1
	2008	6,525.0	36,609.3	43,134.3
	2009	6,654.2	38,869.7	45,523.8
	2010	6,079.2	41,343.6	47,422.8
	2011	5,450.4	50,714.3	56,164.7
	2012	5,122.8	42,796.8	47,919.6
	2013	4,703.5	34,857.4	39,560.9
	2014	4,300.4	33,850.4	38,150.8
	2015	4,176.6	35,080.0	39,256.6
	2016	3,748.0	38,619.8	42,367.8
W	1996	4,363.4	27,294.2	31,657.6
	1997	4,475.2	24,776.4	29,251.6
	1998	4,335.4	26,082.7	30,418.1
	1999	4,043.1	27,457.6	31,500.6
	2000	3,940.7	26,203.8	30,144.5
	2001	4,334.4	24,678.6	29,012.9
	2002	4,032.0	32,151.3	36,183.3
	2003	4,189.3	32,931.0	37,120.3
	2004	4,180.1	32,055.8	36,236.0
	2005	4,194.1	26,184.2	30,378.3
	2006	4,357.4	23,424.4	27,781.8
	2007	4,419.4	25,293.3	29,712.8
	2008	4,231.3	27,020.6	31,251.9
	2009	4,072.4	26,891.5	30,963.9
	2010	4,090.7	29,556.3	33,647.0
	2011	3,820.1	41,902.9	45,723.0
	2012	3,184.8	33,147.1	36,332.0
	2013	3,421.8	26,917.2	30,339.0
	2014	3,137.9	25,177.2	28,315.1
	2015	3,244.0	26,311.6	29,555.6
	2016	2,652.9	28,177.2	30,830.1
Sum	1996	9,633.6	59,786.1	69,419.7
	1997	10,113.8	55,144.0	65,257.8
	1998	10,093.2	59,459.5	69,552.7
	1999	9,763.0	61,121.3	70,884.3
	2000	9,937.7	58,489.2	68,426.9
	2001	10,430.8	57,494.4	67,925.2
	2002	10,496.9	69,325.8	79,822.7
	2003	10,047.9	74,491.4	84,539.3
	2004	9,929.3	72,692.5	82,621.9
	2005	10,273.1	61,219.9	71,493.0
	2006	10,868.3	56,792.2	67,660.4
	2007	11,072.3	59,397.5	70,469.8
	2008	10,756.3	63,629.9	74,386.2
	2009	10,726.6	65,761.1	76,487.7
	2010	10,169.9	70,899.9	81,069.8
	2011	9,270.5	92,617.2	101,887.7
	2012	8,307.6	75,943.9	84,251.5
	2013	8,125.3	61,774.6	69,899.9
	2014	7,438.3	59,027.6	66,465.9
	2015	7,420.6	61,391.5	68,812.2
	2016	6,400.9	66,797.0	73,197.9

We can also graph these figures by type and sex:

```
> par( mar=c(3,4,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> fYLLi <- function(){
+ plot( NA,
+       xlim=c(1996.5,2017),
+       ylim=c(0,51), yaxs="i",
+       xlab="Year of diabetes diagnosis",
+       ylab="" )
+ mtext( "Future life years lost (1000s) to diabetes",
+         side=2, line=2, las=0, cex=1-0.17*(all(par('mfcol')==c(2,2))) )
+ abline( h=0:11*5, col=gray(0.9) )
+ axis( side=1, at=1996:2017, tcl=-0.3, labels=NA )
+ axis( side=2, at=0:10*5, tcl=-0.3, labels=NA )
+ matlines( as.numeric(dimnames(YLLic)[[2]])+0.5,
+           cbind( YLLic["M",,"T1D","Sum"],
+                  YLLic["M",,"T2D","Sum"],
+                  YLLic["W",,"T1D","Sum"],
+                  YLLic["W",,"T2D","Sum"] ) / 1000,
+          col=rep(c("blue","red"),each=2), lty=c("21","solid"),
+          lend="butt", lwd=4, type="l" )
+ }
> fYLLi()
```

We can also compute the *average* number of years lost among the diagnosed patients, so the burden *per person* diagnosed a given year.

```
> fCtable( YLLic.p <- YLLic[,,,5] / apply( ic, 1:3, sum ),
+          w=5, d=1, col.vars=c(3,1) )
     S.in   T1D          T2D
     sex      M     W     M     W
date
1996         7.8   8.5   5.3   5.2
1997         8.2   9.2   5.2   5.0
1998         8.8   9.5   5.1   4.9
1999         9.7   9.8   5.0   4.8
2000        10.1  10.1   4.9   4.7
2001        10.4  10.4   4.8   4.5
2002        10.7  10.4   4.6   4.4
2003        10.7  10.9   4.5   4.3
2004        11.3  10.8   4.4   4.1
2005        11.8  11.1   4.3   4.0
2006        11.8  11.4   4.1   3.9
2007        11.8  11.5   3.9   3.7
2008        12.0  11.5   3.7   3.6
2009        11.7  11.4   3.6   3.5
2010        11.5  11.1   3.5   3.4
2011        11.0  10.7   3.3   3.2
2012        10.5  10.1   3.3   3.3
2013        10.0   9.7   3.4   3.4
2014         9.2   9.2   3.4   3.4
2015         8.8   8.7   3.5   3.4
2016         8.1   8.4   3.6   3.6
> str( YLLic.p )
 num [1:2, 1:21, 1:2] 7.77 8.46 8.24 9.15 8.76 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:21] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
```

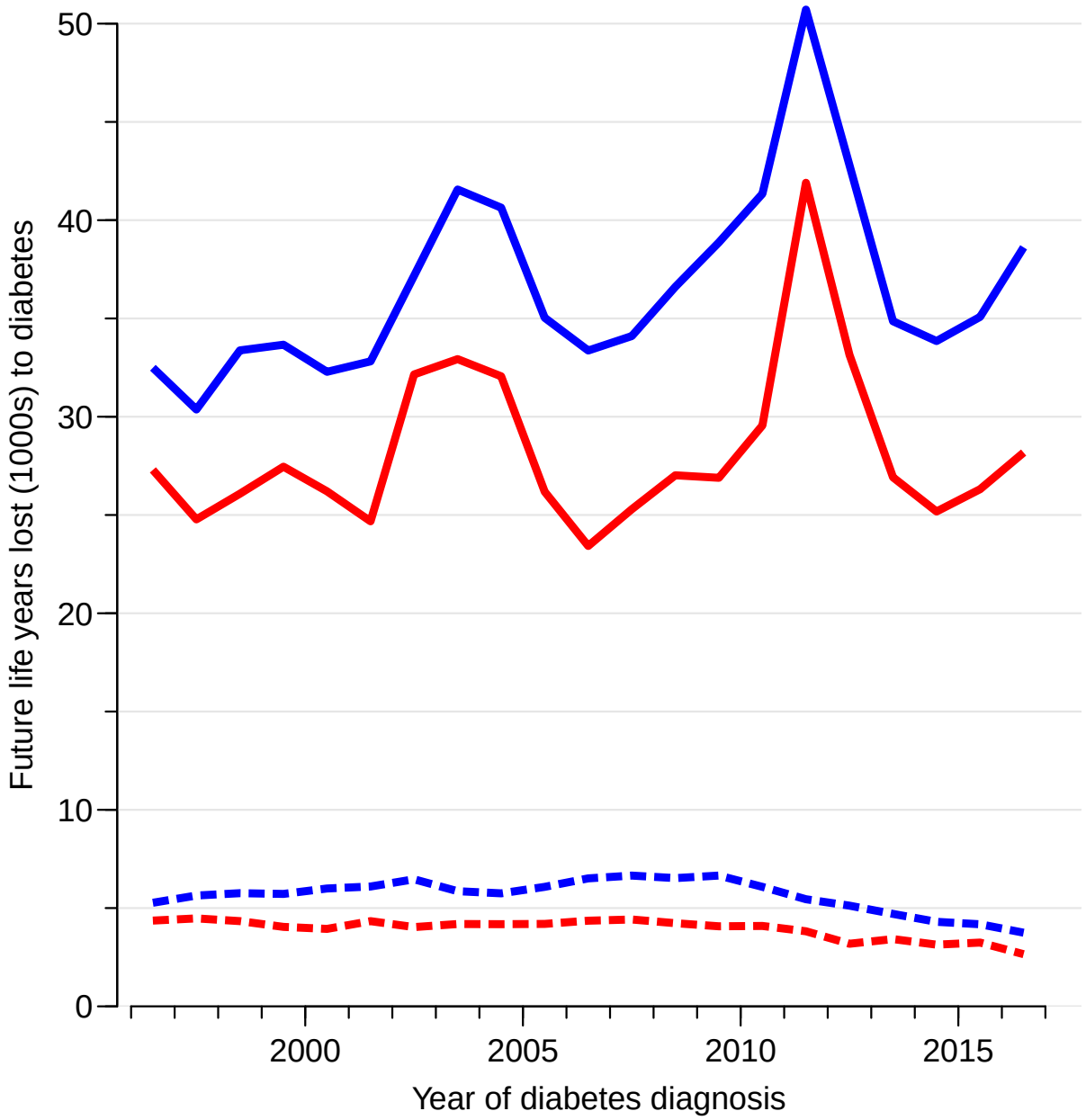


Figure 12.20: *Future years of life lost to diabetes in patients* newly *diagnosed each year 1996–2016. Blue lines are men, red women, dotted lines are T1D and full lines T2D.* `./graph/yll-cod-YLLic`

. . . and make a plot of these average numbers:

```
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> fYLLipp <- function(){
+ plot( NA,
+       xlim=c(1996.5,2017),
+       ylim=c(0,13), yaxs="i",
+       xlab="Year of diabetes diagnosis",
+       ylab="Years lost to diabetes per patient")
+ abline( h=0:13, col=gray(0.9) )
+ axis( side=1, at=1996:2017, tcl=-0.3, labels=NA )
```

```
+ axis( side=2, at=0:13, tcl=-0.3, labels=NA )
+ matlines( as.numeric(dimnames(YLLic.p)[[2]])+0.5,
+           cbind( YLLic.p["M",,"T1D"],
+                  YLLic.p["M",,"T2D"],
+                  YLLic.p["W",,"T1D"],
+                  YLLic.p["W",,"T2D"] ),
+          col=rep(c("blue","red"),each=2), lty=c("21","solid"),
+          lend="butt", lwd=4, type="l" )
+ }
> fYLLipp()
```

We see that the patterns for incident and prevalent cases is very much the same, owing to the fact that the age-distributionas are quite similar.

```
> # utility to plot thins in corners
> cnr <- function( x=4, y=97 ) { xy <- par("usr")
+ list( x = xy[1]*(1-x/100)+xy[2]*x/100,
+       y = xy[3]*(1-y/100)+xy[4]*y/100 ) }
> par( mfrow=c(2,2), mar=c(3,3,1,1), oma=c(0,1,0,0), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> fYLL()     ; text( cnr(), "a", font=2, adj=c(0,1) )
> fYLLpp()   ; text( cnr(), "b", font=2, adj=c(0,1) )
> fYLLi()    ; text( cnr(), "c", font=2, adj=c(0,1) )
> fYLLipp()  ; text( cnr(), "d", font=2, adj=c(0,1) )


> par( mfcol=c(2,2), mar=c(3,3,1,1), oma=c(0,1,0,0), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> fYLL()     ; text( cnr(), "a", font=2, adj=c(0,1) )
> fYLLpp()   ; text( cnr(), "c", font=2, adj=c(0,1) )
> fYLLi()    ; text( cnr(), "b", font=2, adj=c(0,1) )
> fYLLipp()  ; text( cnr(), "d", font=2, adj=c(0,1) )


----------------------
2019-05-27 at 15:44:08
Time elapsed: 00:02:07
----------------------
```

. . . now input from `ltr.tex`

12.8 Life time risk

```
> library(Epi)
> start()

--------------------------------------
Home: E:/workdata/705093/BXC/demoDM/nyr
Time: 2019-03-27 16:50:13
--------------------------------------

> load( file="../nydata/trmods.Rda" )
> load( file="../nydata/inits.Rda" )
> str( T1cvd$M$data )
```

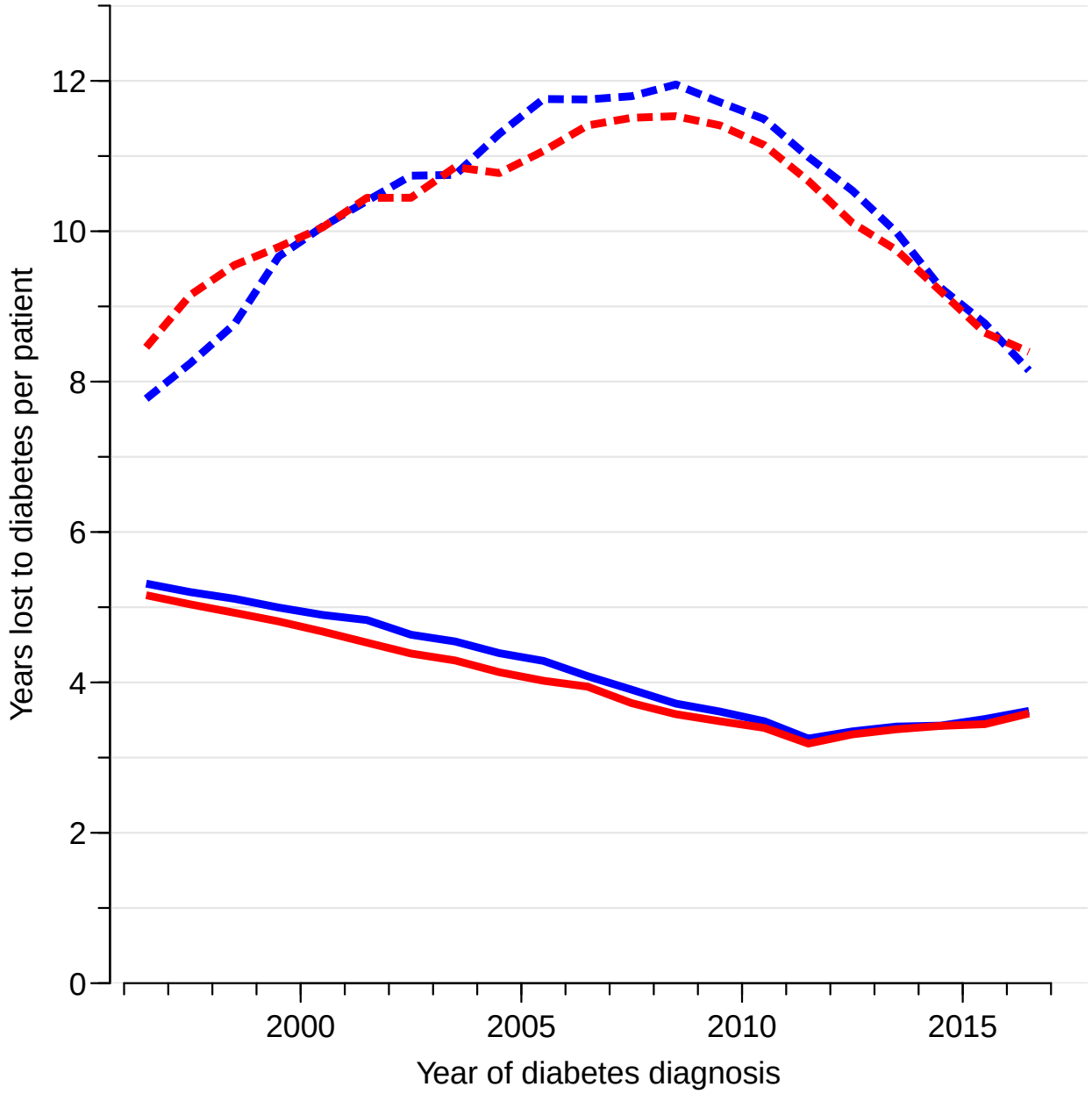


Figure 12.21: *Future years of life lost to diabetes* per patient *diagnosed each year 1996–2016. Blue lines are men, red women, dotted lines are T1D and full lines T2D. This is an average of the years of life lost by age, averaged by the age-distribution of diagnosed patients. T1D is much higher than T2D because the persons diagnosed with T1D are much younger.* `./graph/yll-cod-YLLic-pp`

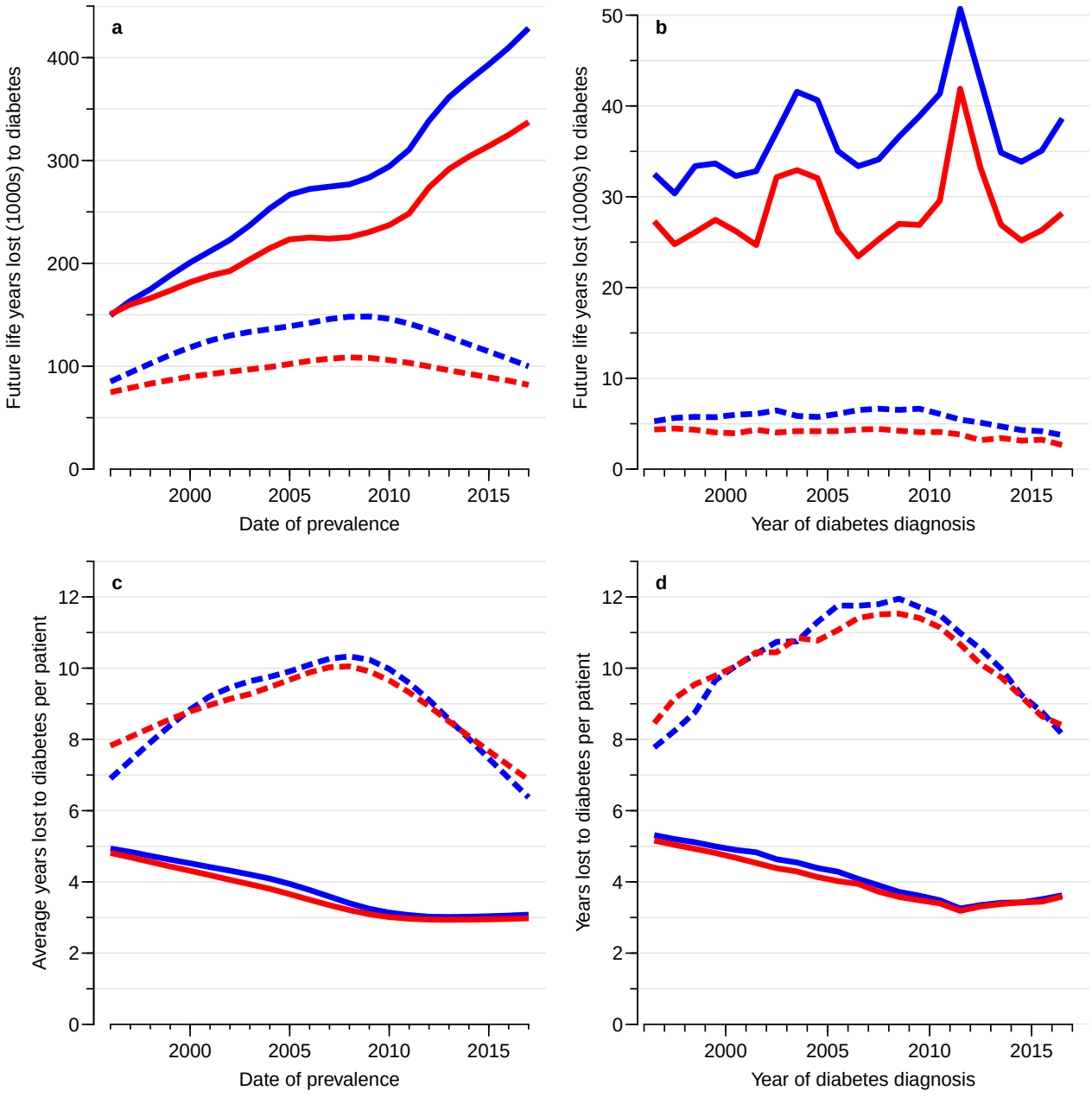


Figure 12.22: *Liftetime lost lost to diabetes in the Danish population:*
a: Total future lifetime lost for persons with diabetes prevalent at each date.
b: Future lifetime lost per person *with diabetes at each date.*
c: Total future lifetime lost for newly diagnosed persons each calendar year.
d: Future lifetime lost per person *for newly diagnosed persons each calendar year.*
`./graph/yll-cod-YLL4`

```
'data.frame':        88397 obs. of  18 variables:
 $ sex  : Factor w/ 2 levels "M","W": 1 1 1 1 1 1 1 1 1 1 ...
 $ state: Factor w/ 3 levels "noDM","T1","T2": 2 2 2 2 2 2 2 2 2 2 ...
 $ A    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ P    : num  1997 1999 1999 2000 2000 ...
 $ C    : num  1996 1998 1999 1999 1999 ...
 $ dur  : num  0.1 0.1 0.1 0.1 0.35 0.1 0.1 0.1 0.35 0.1 ...
 $ Dcvd : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dcan : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Dres : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Doth : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y    : num  1.43e-04 2.26e-05 7.67e-05 3.79e-04 1.76e-04 ...
 $ T1   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ T2   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ D    : num  0 0 0 0 0 0 0 0 0 0 ...
 $ Ax   : num  0.667 0.667 0.333 0.667 0.667 ...
 $ Px   : num  1997 1999 2000 2000 2000 ...
 $ Cx   : num  1997 1999 1999 2000 2000 ...
 $ gP   : Factor w/ 7 levels "1996-1998","1999-2001",..: 1 2 2 2 2 2 3 3 3 3 ...
```

In order to assess the lifetime risk of T1D and T2D we need a slightly different set up of the states of disease and death than the one used when evaluating the years of life lost; instead of death states classified by cause of death we need death states classified by disease state *before* death, that is as seen in figure **??**:

```
> states <- levels( T1cvd$M$data$state )
> ( states <- c( states, paste0("D-", states ) ) )
[1] "noDM"   "T1"     "T2"     "D-noDM" "D-T1"   "D-T2"
> Tm <- matrix(NA,6,6)
> rownames(Tm) <-
+ colnames(Tm) <- states
> Tm["noDM",c("T1","T2","D-noDM")] <-
+ Tm["T1","D-T1"] <-
+ Tm["T2","D-T2"] <- 1
> boxes.matrix( Tm, boxpos=list(x=rep(c(10,90),each=3),
+                               y=rep(c(50,90,10),2)),
+               hmult=3 )
```

12.9 Transition probabilities for life time risk

We now set up state-transition matrices for each sex and date (the date where we take cross-sectional rates to compute state occupancy probabilities), one per age, all stored in an array. The last two dimensions of the array form the transition matrix, at a given time (age, the third last dimension):

```
> int <- 1/12
> a.pt <- seq(int,100,int) - int/2
> states <- rownames( Tm )
> Tr <- ZArray( list( sex = levels(T1cvd$M$data$sex),
+                     date = 1996:2017,
+                      age = a.pt,
+                     from = states,
+                       to = states ) )
> fC( length(Tr) ) ; str( Tr )
```

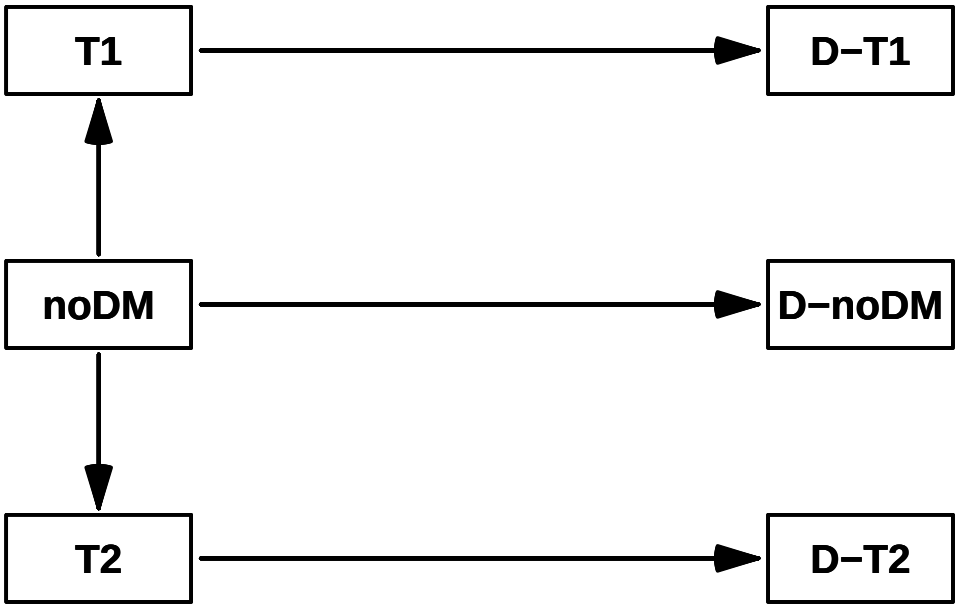


Figure 12.23: *States used in the calculation of life-time risk of T1D and T2D. Note that the rates are identical to (or simple sums of) rates used in the setup shown in figure 12.1.*
`./graph/ltr-boxltr`

```
[1] "1,900,800"
 num [1:2, 1:22, 1:1200, 1:6, 1:6] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ age : chr [1:1200] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667"
  ..$ from: chr [1:6] "noDM" "T1" "T2" "D-noDM" ...
  ..$ to  : chr [1:6] "noDM" "T1" "T2" "D-noDM" ...
```

So we now fill the transition probabilities in; recall that the variable `Y` in the dataset `rt` is measured in millenia, so when we make predictions for 1-month intervals we must do this for a value of Y equal to 1/12000 (that is `int`/1000).

We fill in the transition probabilities by just using the cumulative intensities over the intervals; when these are so small as in this case they are extremely close to the transition probabilities. Furthermore we use the sum of the cause-specific mortalities to represent the

```
> system.time(
+ for( sx in dimnames(Tr)[[1]] )
+ for( dd in dimnames(Tr)[[2]] )
+    {
+    nd <- data.frame( Px = as.numeric(dd),
+                      Ax = a.pt,
+                       Y = int/1000 )
+    Tr[sx,dd,,"noDM","T1"    ] <- ci.pred(  nDT1[[sx]], nd )[,1]
+    Tr[sx,dd,,"noDM","T2"    ] <- ci.pred(  nDT2[[sx]], nd )[,1]
+
+    Tr[sx,dd,,"noDM","D-noDM"] <- ci.pred( nDcvd[[sx]], nd )[,1]+
+                                  ci.pred( nDcan[[sx]], nd )[,1]+
```

```
+                                  ci.pred( nDres[[sx]], nd )[,1]+
+                                  ci.pred( nDoth[[sx]], nd )[,1]
+
+    Tr[sx,dd,,  "T1","D-T1"  ] <- ci.pred( T1cvd[[sx]], nd )[,1]+
+                                  ci.pred( T1can[[sx]], nd )[,1]+
+                                  ci.pred( T1res[[sx]], nd )[,1]+
+                                  ci.pred( T1oth[[sx]], nd )[,1]
+
+    Tr[sx,dd,,  "T2","D-T2"  ] <- ci.pred( T2cvd[[sx]], nd )[,1]+
+                                  ci.pred( T2can[[sx]], nd )[,1]+
+                                  ci.pred( T2res[[sx]], nd )[,1]+
+                                  ci.pred( T2oth[[sx]], nd )[,1]
+    } )
   user  system elapsed
   7.03    1.47    8.50
> Tdiag <- 1 - apply( Tr, 1:4, sum )
> for( i in 1:dim(Tr)[5] ) Tr[,,,i,i] <- Tdiag[,,,i]
> range( Tr )
[1] 0 1
```

We can inspect three instances of the transition probability matrices; we multiply by 10^6 to get readable numbers (remember these are 1-month transition probabilities).

```
> fCtable( addmargins(Tr["M","2016",c(30,50,70)*12,,]*10^6,3),
+                 row.vars=1:2 )
                        to      noDM        T1        T2    D-noDM      D-T1      D-T2       S
age              from
29.9583333333333 noDM        999,899        19        42        40         .         . 1,000,0
                 T1                .   999,838         .         .       162         . 1,000,0
                 T2                .         .   999,802         .         .       198 1,000,0
                 D-noDM            .         .         . 1,000,000         .         . 1,000,0
                 D-T1              .         .         .         . 1,000,000         . 1,000,0
                 D-T2              .         .         .         .         . 1,000,000 1,000,0
49.9583333333333 noDM        999,341         8       431       220         .         . 1,000,0
                 T1                .   999,226         .         .       774         . 1,000,0
                 T2                .         .   999,326         .         .       674 1,000,0
                 D-noDM            .         .         . 1,000,000         .         . 1,000,0
                 D-T1              .         .         .         . 1,000,000         . 1,000,0
                 D-T2              .         .         .         .         . 1,000,000 1,000,0
69.9583333333333 noDM        997,533         4       958     1,506         .         . 1,000,0
                 T1                .   996,429         .         .     3,571         . 1,000,0
                 T2                .         .   997,238         .         .     2,762 1,000,0
                 D-noDM            .         .         . 1,000,000         .         . 1,000,0
                 D-T1              .         .         .         . 1,000,000         . 1,000,0
                 D-T2              .         .         .         .         . 1,000,000 1,000,0
```

12.10 State occupancy probabilities

We can now compute state occupancy probabilities from various starting points; we only use noDM as starting state but starting ages for calculation of cumulative risk as well as expected sojourn time in states. Note that we are computing the state occupancy probabilities at the borders of the age-intervals, using starting age 0.

```
> Pr <- ZArray( c( dimnames(Tr)[1:2],
+                  list( age = c(0,a.pt+int/2), # border of age intervals
+                      state = states ) ) )
> str( Pr )
 num [1:2, 1:22, 1:1201, 1:6] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 4
  ..$ sex  : chr [1:2] "M" "W"
  ..$ date : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ age  : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ state: chr [1:6] "noDM" "T1" "T2" "D-noDM" ...
```

Now it only remains to fill values into the object:

```
> system.time(
+ for( sx in dimnames(Pr)[[1]] ) # sex
+ for( dd in dimnames(Pr)[[2]] ) # date.in
+    {
+    # for ages before the probability is 1 for the initial state
+    Pr[sx,dd,1,"noDM"] <- 1
+    # and transitions occur at the remaining ages
+    for( ca in 2:dim(Pr)[3] )
+       Pr[sx,dd,ca,] <- Pr[sx,dd,ca-1,] %*% Tr[sx,dd,ca-1,,]
+    } )
   user  system elapsed
   0.78    0.00    0.78
```

We check sanity of resulting array:

```
> range( Pr ) ; range( apply( Pr, 1:3, sum ) )
[1] 0 1
[1] 1 1
```

Thus we now have the state occupancy probabilities for all 6 states by age, each of these classified by the conditioning variables (entry characteristics): sex and calendar time.

To illustrate how this pans out over life we plot the state probabilities:

```
> xr <- function(x)       c(x,rev(x))
> yr <- function(x,y=0*x) c(x,rev(y))
> clr <- c("forestgreen","orange","red")
> clr <- c( clr, adjustcolor( rev(clr), alpha.f=0.5 ) )
> str( Pr )
 num [1:2, 1:22, 1:1201, 1:6] 1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, "dimnames")=List of 4
  ..$ sex  : chr [1:2] "M" "W"
  ..$ date : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ age  : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ state: chr [1:6] "noDM" "T1" "T2" "D-noDM" ...
> Px <- Pr[,,,c(1,1,3,2,5,6,4)]
> Px[,,,1] <- 0
> Px <- aperm( apply( Px, 1:3, cumsum ), c(2:4,1) )
> aa <- as.numeric(dimnames(Pr)[['age']])
> pp <- "2011"
> sx <- "M"
> a1 <- function(){
+       axis( side=1, at=0:10*10 )
+       axis( side=1, at=0:20*5, labels=NA, tcl=-0.3 )
```

```
+       }
> a2 <- function(s,n=TRUE){
+       axis( side=s, at=0:10/10, if(!n) labels=NA )
+       axis( side=s, at=0:20/20, labels=NA, tcl=-0.3 )
+       axis( side=s, at=0:100/100, labels=NA, tcl=-0.15 )
+       }
> pl <- function( sx, pp, yl=c(50,100) )
+       {
+ plot( NA, xlim=yl,  ylim=c(0,1),
+           xaxs="i", yaxs="i",
+           xaxt="n", yaxt="n",
+           xlab="",  ylab="" )
+ for( i in 1:6 )
+ polygon( xr(aa), yr(Px[sx,pp,,i],Px[sx,pp,,i+1]),
+          col=clr[i], border="transparent" )
+ lines( aa, Px[sx,pp,,4] )
+       }
> yrs <- paste(2006+0:2*5)
> par( mfrow=c(3,2), mar=c(1,1.5,1,1.5), oma=c(3,2,2,2),
+      mgp=c(3,1,0)/1.6, bty="n", las=1 )
> for( yy in yrs )
+    {
+ pl("M",yy)           ; a1() ; a2(2,T) ; a2(4,F)
+ pl("W",yy,c(100,50)) ; a1() ; a2(2,T) ; a2(4,T)
+    }
> mtext( c("Men","Women"), at=c(1,3)/4, side=3, outer=TRUE, cex=0.66 )
> mtext( yrs, at=c(5,3,1)/6, side=2, outer=TRUE, line=1, las=0, cex=0.66 )
> mtext( c("Age"), at=c(1,3)/4, side=1, outer=TRUE, line=1.5, cex=0.66 )
```

12.11 Cumulative risk of T2D

```
> str( Pr )
 num [1:2, 1:22, 1:1201, 1:6] 1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, "dimnames")=List of 4
  ..$ sex  : chr [1:2] "M" "W"
  ..$ date : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ age  : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
  ..$ state: chr [1:6] "noDM" "T1" "T2" "D-noDM" ...
> cr2 <- Pr[,,,"T2"] +  Pr[,,,"D-T2"]
> str(cr2)
 num [1:2, 1:22, 1:1201] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ age : chr [1:1201] "0" "0.0833333333333333" "0.166666666666667" "0.25" ...
> aa <- as.numeric(dimnames(Pr)[['age']])
> pp <- c(2000,2004,2008,2012,2016)
> a1 <- function(){
+       axis( side=1, at=0:10*10 )
+       axis( side=1, at=0:20*5, labels=NA, tcl=-0.3 )
+       }
> a2 <- function(s=2,n=TRUE){
+       axis( side=s, at=0:10*10, if(!n) labels=NA )
```

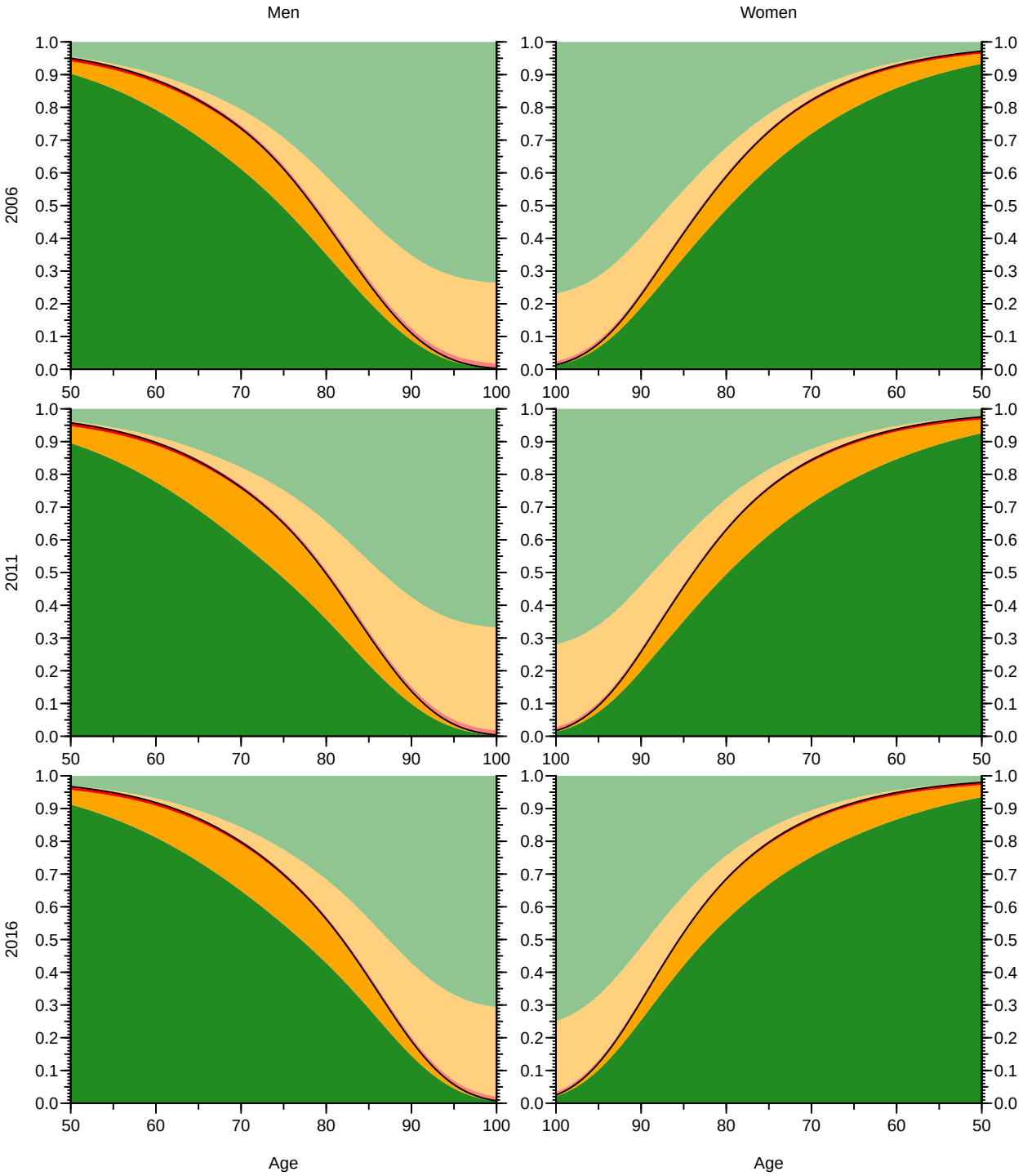


Figure 12.24: *Probability of being in the state noDM (green), T2D (orange), T1D (red), or dead from each of these states (pale colors).* `./graph/ltr-st-occ`

```
+       axis( side=s, at=0:20*5, labels=NA, tcl=-0.3 )
+       axis( side=s, at=0:100, labels=NA, tcl=-0.15 )
+       }
> par( mfrow=c(2,1), mar=c(3,1,1,1), oma=c(2,2,0,2), mgp=c(3,1,0)/1.6 )
> for( sx in c("M","W") ){
```

`ltr.tex`

```
+ cl <- ifelse(sx=="M","blue","red")
+ plot( NA, xlim=c(0,100), ylim=c(0,33),
+           xaxs="i", yaxs="i",
+           xaxt="n", yaxt="n",
+           xlab="",  ylab="" )
+ matlines( aa, t(cr2[sx,paste(pp),]*100),
+           lty=1, lwd=3, col=cl )
+ a1() ; a2(2) ; a2(4)
+ text( 5,30, sx, col=cl, font=2, cex=2 )
+ text( 95, cr2[sx,paste(pp),1200]*100+c(1,1,1,1,-7)/7, paste(pp), adj=c(1,0), col=cl )
+ abline(v=45,h=cr2[sx,"2016",12*45+1]*100)
+ }
> mtext( "Cumulative risk of T2D (%)", line=1, side=4, outer=TRUE, las=0 )
```

12.12 Lifetime risk of diabetes

As an alternative to this we can now show the cumulative risks of T1D resp. T2D (and their sum) and how this has evolved over time:

```
> round( ftable( Pr[,,c("50","100"),], row.vars=c(3,2) )*100, 1 )
```

	sex	M						W					
	state	noDM	T1	T2	D-noDM	D-T1	D-T2	noDM	T1	T2	D-noDM	D-T1	D-T2
age	date												
50	1996	89.4	1.1	2.5	6.6	0.2	0.3	93.3	0.8	1.9	3.7	0.1	0.1
	1997	89.7	1.0	2.5	6.3	0.2	0.3	93.4	0.8	1.9	3.6	0.1	0.1
	1998	89.9	1.0	2.6	6.1	0.2	0.2	93.5	0.8	2.0	3.5	0.1	0.1
	1999	90.1	1.0	2.7	5.8	0.2	0.2	93.6	0.8	2.1	3.3	0.1	0.1
	2000	90.3	0.9	2.8	5.6	0.2	0.2	93.6	0.7	2.2	3.2	0.1	0.1
	2001	90.4	0.9	2.9	5.4	0.2	0.2	93.7	0.7	2.3	3.1	0.1	0.1
	2002	90.5	0.9	3.0	5.1	0.2	0.2	93.7	0.7	2.4	3.0	0.1	0.1
	2003	90.6	0.9	3.1	4.9	0.2	0.2	93.7	0.7	2.5	2.9	0.1	0.1
	2004	90.6	0.9	3.3	4.8	0.2	0.2	93.6	0.7	2.7	2.8	0.1	0.1
	2005	90.5	1.0	3.5	4.6	0.2	0.2	93.5	0.7	2.8	2.7	0.1	0.1
	2006	90.3	1.0	3.7	4.5	0.2	0.3	93.4	0.8	3.1	2.6	0.1	0.1
	2007	90.1	1.0	4.0	4.4	0.2	0.3	93.2	0.8	3.3	2.5	0.1	0.2
	2008	89.9	1.0	4.4	4.3	0.2	0.3	92.9	0.8	3.6	2.4	0.1	0.2
	2009	89.6	1.0	4.7	4.1	0.2	0.3	92.8	0.8	3.9	2.3	0.1	0.2
	2010	89.5	1.0	5.0	4.0	0.2	0.3	92.6	0.8	4.1	2.3	0.1	0.2
	2011	89.6	1.0	5.2	3.8	0.2	0.3	92.6	0.8	4.2	2.2	0.1	0.2
	2012	89.8	1.0	5.2	3.6	0.2	0.3	92.7	0.8	4.2	2.1	0.1	0.2
	2013	90.1	1.0	5.1	3.4	0.1	0.3	92.8	0.8	4.1	2.0	0.1	0.1
	2014	90.5	1.0	4.9	3.3	0.1	0.2	93.1	0.8	4.0	2.0	0.1	0.1
	2015	90.9	1.0	4.7	3.1	0.1	0.2	93.3	0.8	3.8	1.9	0.1	0.1
	2016	91.2	1.0	4.5	3.0	0.1	0.2	93.5	0.8	3.7	1.8	0.1	0.1
	2017	91.6	1.0	4.3	2.8	0.1	0.2	93.7	0.8	3.5	1.7	0.1	0.1
100	1996	0.1	0.0	0.0	80.4	2.0	17.5	0.8	0.0	0.1	82.1	1.5	15.5
	1997	0.1	0.0	0.0	80.0	1.9	18.0	0.8	0.0	0.1	81.9	1.4	15.8
	1998	0.1	0.0	0.0	79.6	1.8	18.5	0.8	0.0	0.1	81.7	1.4	16.1
	1999	0.2	0.0	0.0	79.2	1.7	19.0	0.9	0.0	0.1	81.4	1.3	16.4
	2000	0.2	0.0	0.0	78.7	1.7	19.5	0.9	0.0	0.1	81.1	1.2	16.8
	2001	0.2	0.0	0.0	78.2	1.6	20.0	0.9	0.0	0.1	80.7	1.2	17.1
	2002	0.2	0.0	0.0	77.6	1.6	20.6	0.9	0.0	0.1	80.2	1.1	17.6
	2003	0.2	0.0	0.0	76.9	1.5	21.3	1.0	0.0	0.1	79.7	1.1	18.1
	2004	0.2	0.0	0.0	76.0	1.5	22.2	1.1	0.0	0.1	78.9	1.1	18.8
	2005	0.2	0.0	0.0	74.9	1.5	23.3	1.1	0.0	0.2	78.0	1.1	19.6

```
     2006        0.3  0.0  0.0   73.5  1.5 24.7  1.2  0.0  0.2   76.9  1.1 20.7
     2007        0.3  0.0  0.0   71.9  1.5 26.3  1.2  0.0  0.2   75.6  1.1 21.9
     2008        0.3  0.0  0.1   70.2  1.5 28.0  1.2  0.0  0.2   74.3  1.1 23.1
     2009        0.3  0.0  0.1   68.6  1.5 29.6  1.3  0.0  0.3   73.1  1.1 24.2
     2010        0.3  0.0  0.1   67.3  1.5 30.8  1.3  0.0  0.3   72.3  1.1 25.1
     2011        0.3  0.0  0.1   66.7  1.4 31.5  1.4  0.0  0.3   71.9  1.0 25.4
     2012        0.3  0.0  0.1   66.9  1.4 31.3  1.5  0.0  0.3   72.1  1.0 25.1
     2013        0.4  0.0  0.1   67.5  1.4 30.6  1.6  0.0  0.3   72.6  1.0 24.4
     2014        0.5  0.0  0.1   68.5  1.3 29.6  1.8  0.0  0.3   73.4  1.0 23.5
     2015        0.5  0.0  0.1   69.6  1.3 28.5  1.9  0.0  0.3   74.2  1.0 22.6
     2016        0.6  0.0  0.1   70.6  1.3 27.4  2.1  0.0  0.4   74.9  1.0 21.6
     2017        0.7  0.0  0.1   71.6  1.2 26.3  2.3  0.0  0.4   75.6  1.0 20.8
> cPr <- addmargins( Pr[,,,2:3] + Pr[,,,5:6], 4 )*100
> round( ftable( cPr[,,c("50","100"),], row.vars=2 ), 1 )
     sex      M                                W
     age     50             100              50             100
     state   T1   T2  Sum   T1   T2  Sum   T1   T2  Sum   T1   T2  Sum
date
1996        1.3  2.7  4.0  2.0 17.5 19.5  0.9  2.0  2.9  1.5 15.5 17.1
1997        1.2  2.8  4.0  1.9 18.0 19.9  0.9  2.1  3.0  1.4 15.8 17.3
1998        1.2  2.8  4.0  1.8 18.5 20.3  0.9  2.1  3.0  1.4 16.1 17.5
1999        1.2  2.9  4.1  1.7 19.0 20.7  0.9  2.2  3.1  1.3 16.5 17.8
2000        1.1  3.0  4.2  1.7 19.5 21.1  0.8  2.3  3.2  1.2 16.8 18.1
2001        1.1  3.1  4.2  1.6 20.0 21.6  0.8  2.4  3.2  1.2 17.2 18.4
2002        1.1  3.2  4.3  1.6 20.6 22.2  0.8  2.5  3.3  1.1 17.7 18.8
2003        1.1  3.3  4.5  1.5 21.3 22.9  0.8  2.6  3.5  1.1 18.2 19.3
2004        1.1  3.5  4.6  1.5 22.2 23.7  0.8  2.8  3.6  1.1 18.9 20.0
2005        1.2  3.7  4.9  1.5 23.3 24.8  0.8  3.0  3.8  1.1 19.8 20.9
2006        1.2  4.0  5.2  1.5 24.7 26.2  0.9  3.2  4.1  1.1 20.9 21.9
2007        1.2  4.3  5.5  1.5 26.3 27.8  0.9  3.5  4.3  1.1 22.1 23.2
2008        1.2  4.7  5.9  1.5 28.1 29.6  0.9  3.8  4.6  1.1 23.4 24.5
2009        1.2  5.0  6.2  1.5 29.7 31.2  0.9  4.0  4.9  1.1 24.5 25.6
2010        1.2  5.3  6.5  1.5 30.9 32.4  0.9  4.2  5.1  1.1 25.3 26.4
2011        1.2  5.4  6.7  1.4 31.5 33.0  0.9  4.3  5.2  1.0 25.7 26.7
2012        1.2  5.4  6.6  1.4 31.4 32.8  0.9  4.3  5.2  1.0 25.4 26.5
2013        1.2  5.3  6.5  1.4 30.7 32.1  0.9  4.3  5.1  1.0 24.8 25.8
2014        1.2  5.1  6.3  1.3 29.7 31.0  0.9  4.1  5.0  1.0 23.9 24.9
2015        1.1  4.9  6.0  1.3 28.6 29.9  0.9  4.0  4.8  1.0 22.9 23.9
2016        1.1  4.7  5.8  1.3 27.5 28.8  0.9  3.8  4.7  1.0 22.0 23.0
2017        1.1  4.5  5.6  1.2 26.4 27.7  0.9  3.7  4.5  1.0 21.1 22.1
> round( ftable( cPr[,,c("50","100"),], col.vars=c(3,1,4) ), 1 )
     age     50                               100
     sex      M                W               M                W
     state   T1   T2  Sum   T1   T2  Sum   T1   T2  Sum   T1   T2  Sum
date
1996        1.3  2.7  4.0  0.9  2.0  2.9  2.0 17.5 19.5  1.5 15.5 17.1
1997        1.2  2.8  4.0  0.9  2.1  3.0  1.9 18.0 19.9  1.4 15.8 17.3
1998        1.2  2.8  4.0  0.9  2.1  3.0  1.8 18.5 20.3  1.4 16.1 17.5
1999        1.2  2.9  4.1  0.9  2.2  3.1  1.7 19.0 20.7  1.3 16.5 17.8
2000        1.1  3.0  4.2  0.8  2.3  3.2  1.7 19.5 21.1  1.2 16.8 18.1
2001        1.1  3.1  4.2  0.8  2.4  3.2  1.6 20.0 21.6  1.2 17.2 18.4
2002        1.1  3.2  4.3  0.8  2.5  3.3  1.6 20.6 22.2  1.1 17.7 18.8
2003        1.1  3.3  4.5  0.8  2.6  3.5  1.5 21.3 22.9  1.1 18.2 19.3
2004        1.1  3.5  4.6  0.8  2.8  3.6  1.5 22.2 23.7  1.1 18.9 20.0
2005        1.2  3.7  4.9  0.8  3.0  3.8  1.5 23.3 24.8  1.1 19.8 20.9
2006        1.2  4.0  5.2  0.9  3.2  4.1  1.5 24.7 26.2  1.1 20.9 21.9
```

```
2007        1.2  4.3  5.5  0.9  3.5  4.3  1.5 26.3 27.8  1.1 22.1 23.2
2008        1.2  4.7  5.9  0.9  3.8  4.6  1.5 28.1 29.6  1.1 23.4 24.5
2009        1.2  5.0  6.2  0.9  4.0  4.9  1.5 29.7 31.2  1.1 24.5 25.6
2010        1.2  5.3  6.5  0.9  4.2  5.1  1.5 30.9 32.4  1.1 25.3 26.4
2011        1.2  5.4  6.7  0.9  4.3  5.2  1.4 31.5 33.0  1.0 25.7 26.7
2012        1.2  5.4  6.6  0.9  4.3  5.2  1.4 31.4 32.8  1.0 25.4 26.5
2013        1.2  5.3  6.5  0.9  4.3  5.1  1.4 30.7 32.1  1.0 24.8 25.8
2014        1.2  5.1  6.3  0.9  4.1  5.0  1.3 29.7 31.0  1.0 23.9 24.9
2015        1.1  4.9  6.0  0.9  4.0  4.8  1.3 28.6 29.9  1.0 22.9 23.9
2016        1.1  4.7  5.8  0.9  3.8  4.7  1.3 27.5 28.8  1.0 22.0 23.0
2017        1.1  4.5  5.6  0.9  3.7  4.5  1.2 26.4 27.7  1.0 21.1 22.1
```

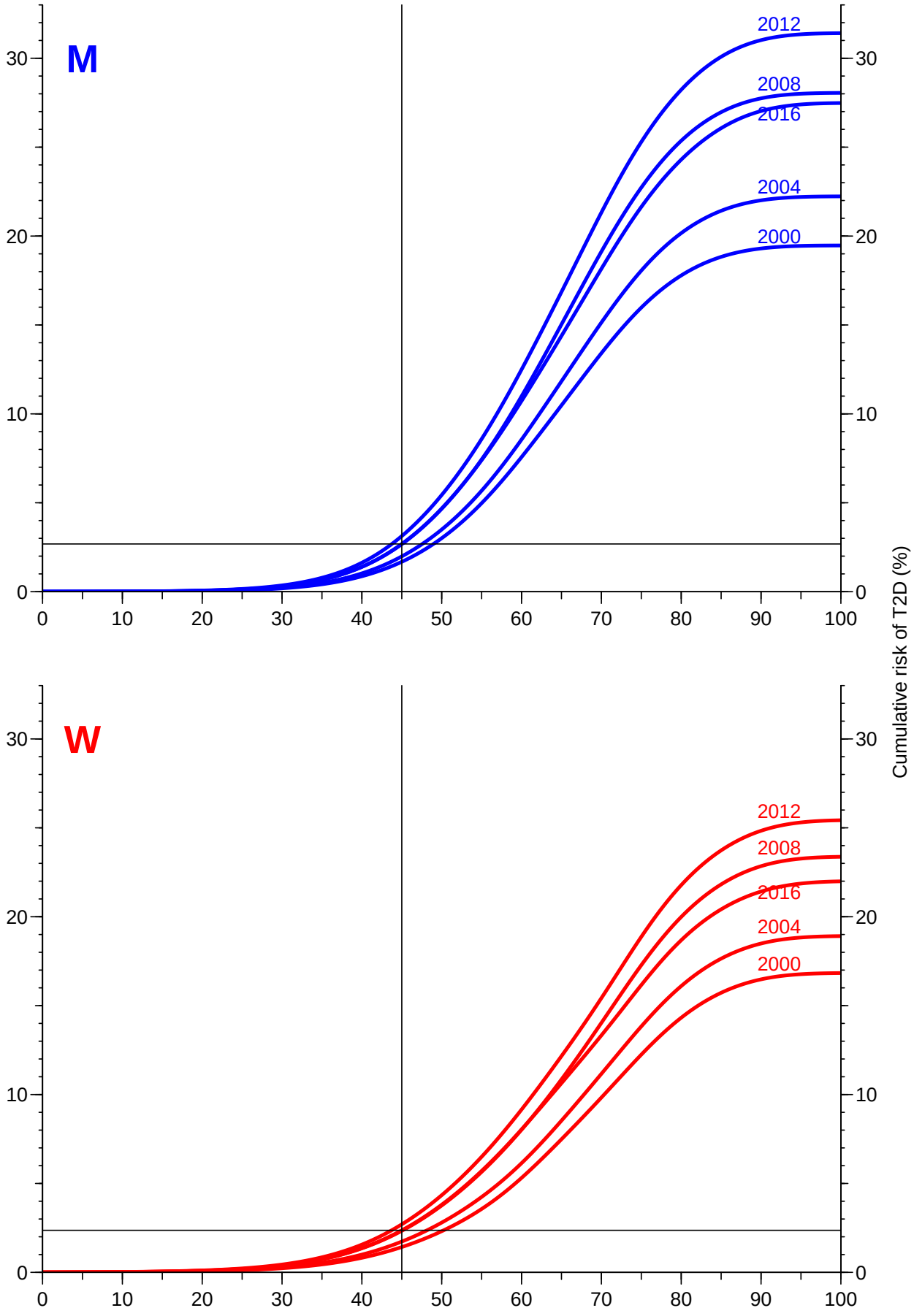
```
> p.pt <- as.numeric(dimnames(cPr)[[2]])
> par( mar=c(0,0,0,0), oma=c(4,4,2,1), mfrow=c(1,2), las=1, bty="n" )
> pllr <- function( sx )
+     {
+ clr <- if( sx=="M" ) "blue" else "red"
+ plot( NA, xlim=c(1996,2017), ylim=c(0,35), yaxs="i", yaxt=if(sx=="M") "s" else "n" )
+ polygon( xr(p.pt), yr(cPr[sx,,"100","T1"]),
+          col=adjustcolor(clr,alpha.f=0.6), border='transparent' )
+ polygon( xr(p.pt), yr(cPr[sx,,"100","T2"]),
+          col=adjustcolor(clr,alpha.f=0.3), border='transparent' )
+ lines( p.pt, cPr[sx,,"100","Sum"], col=clr, lwd=3 )
+ abline( h=1:35, col=gray(0.9), lty="13", lend="butt" )
+ axis( side=1, at=1996:2017, tcl=-0.3, labels=NA )
+ if( sx=="M" ) axis( side=2, at=0:35, tcl=-0.3, labels=NA )
+     }
> pllr("M")
> pllr("W")
> mtext( "Date", side=1, line=2.5, outer=TRUE, las=0 )
> mtext( "Lifetime risk (%)", side=2, line=2.5, outer=TRUE, las=0 )
```

```
> p.pt <- as.numeric(dimnames(cPr)[[2]])
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, las=1, bty="n" )
> clr <- c("blue","red")
> plot( NA, xlim=c(1996,2017), ylim=c(0,32), yaxs="i",
+           xlab="Date", ylab="Lifetime risk (%)" )
> abline( h=1:32, col=gray(0.8), lty="22", lend="butt" )
> matlines( p.pt, t(cPr[,,"100","T1"]), lwd=3, lty="21", lend="butt", col=clr )
> matlines( p.pt, t(cPr[,,"100","T2"]), lwd=3, lty="solid", col=clr )
> axis( side=1, at=1996:2017, tcl=-0.3, labels=NA )
> axis( side=2, at=0:32, tcl=-0.3, labels=NA )
```

12.12.1 Sojourn times

From the array with state occupancy probabilities we can compute the expected sojourn times without DM, with T1D and with T2D as simple integrals of the probabilities. Since the array `Pr` contains the probabilities we can easily compute the expected time spent without diabetes and with T1D resp. T2D.

```
> trz <- function( y ) sum( y[-1]-diff(y)/2 )
> Sj <- apply( Pr[,,,1:3], c(1,2,4), trz ) * int
> round( ftable( addmargins(Sj,3), row.vars=2 ), 1 )
```



`ltr.tex`

Figure 12.25: *Cumulative risk of T2D for men and women, at selected dates.* `./graph/ltr-cr2`

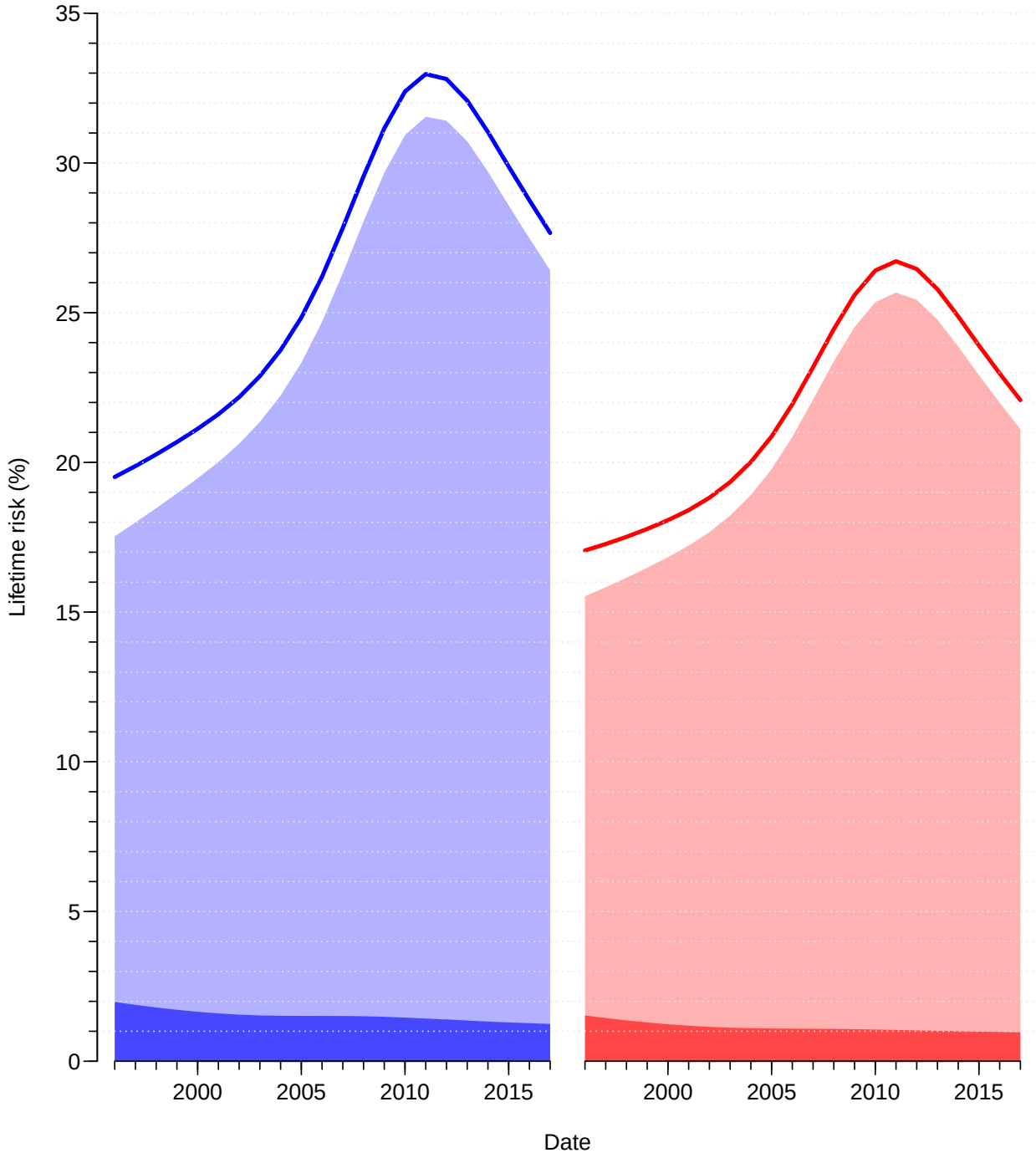


Figure 12.26: *Lifetime risk of diabetes in Denmark at different calendar times. The dense area at the bottom is the lifetime risk of T1D, the total shaded area is the lifetime risk of T2D, and the line at the top shows the lifetime risk of any of the two types of diabetes. Blue is men, red is women.* `./graph/ltr-cumrisk`

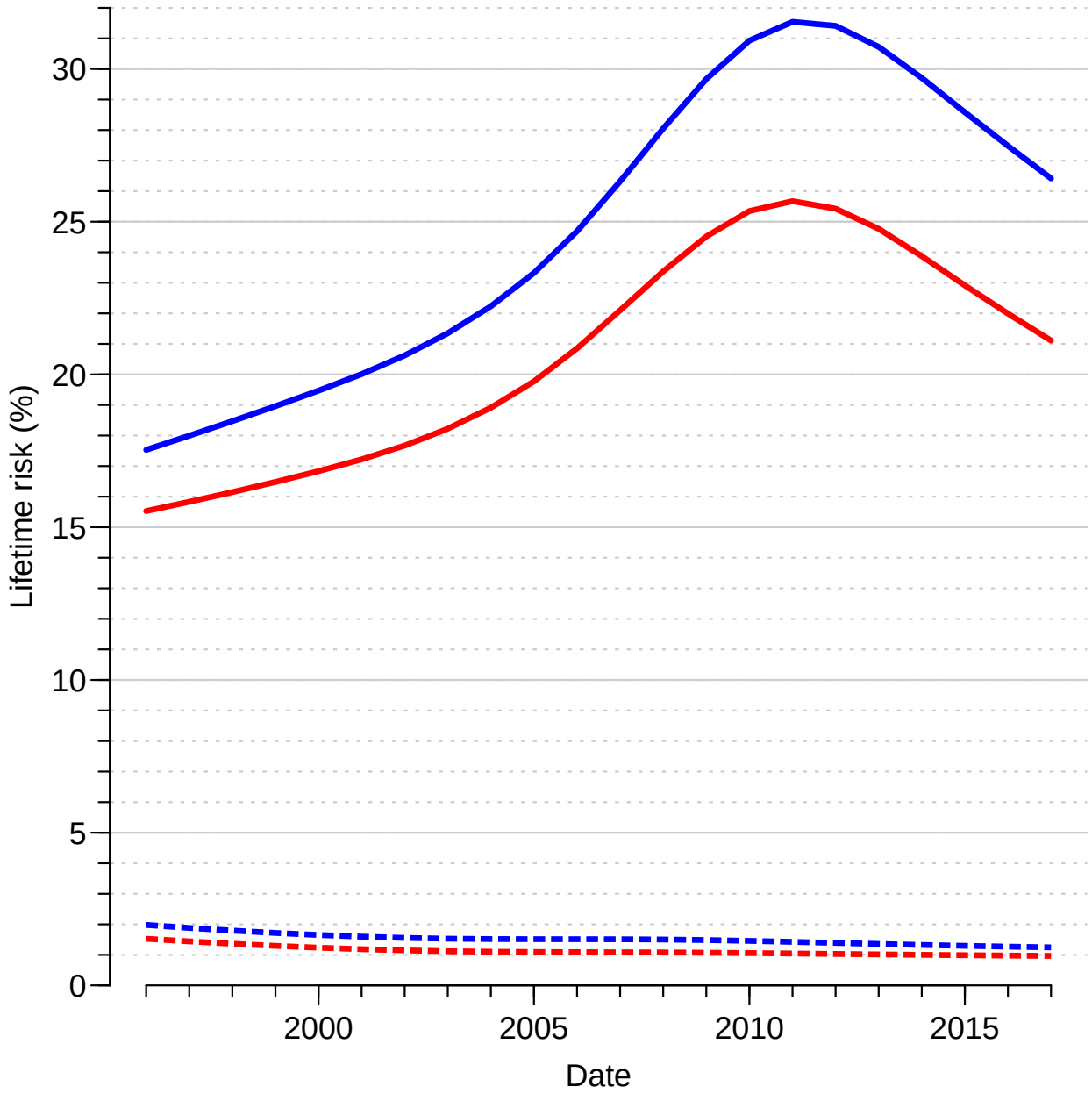


Figure 12.27: *Lifetime risk of diabetes in Denmark at different calendar times. The broken lines are lifetime risk of T1D, the full lines is lifetime risk of T2D. Blue is men, red is women.*
`./graph/ltr-cumrline`

```
     sex      M                    W
     state noDM   T1   T2  Sum noDM   T1   T2  Sum
date
1996       70.2  0.6  2.1 72.8 75.5  0.5  2.0 78.1
1997       70.4  0.5  2.2 73.1 75.7  0.5  2.1 78.3
1998       70.7  0.5  2.3 73.5 75.8  0.5  2.2 78.4
1999       70.9  0.5  2.4 73.8 75.9  0.4  2.3 78.6
2000       71.1  0.5  2.5 74.1 76.0  0.4  2.4 78.8
2001       71.3  0.5  2.6 74.4 76.1  0.4  2.5 79.1
2002       71.5  0.5  2.8 74.7 76.3  0.4  2.6 79.3
2003       71.6  0.5  2.9 75.0 76.4  0.4  2.8 79.6
2004       71.7  0.5  3.1 75.4 76.4  0.4  3.0 79.8
2005       71.8  0.5  3.4 75.6 76.5  0.4  3.2 80.1
2006       71.7  0.5  3.6 75.9 76.4  0.4  3.4 80.3
2007       71.6  0.5  4.0 76.1 76.3  0.4  3.7 80.5
2008       71.4  0.5  4.4 76.3 76.2  0.4  4.0 80.7
2009       71.3  0.5  4.7 76.6 76.2  0.5  4.3 80.9
2010       71.3  0.5  5.0 76.9 76.2  0.5  4.5 81.2
2011       71.4  0.6  5.2 77.2 76.3  0.5  4.7 81.4
2012       71.7  0.6  5.3 77.6 76.6  0.5  4.7 81.7
2013       72.2  0.6  5.2 78.0 76.9  0.5  4.6 82.0
2014       72.7  0.6  5.1 78.4 77.4  0.5  4.5 82.4
2015       73.3  0.6  4.9 78.8 77.8  0.5  4.3 82.7
2016       73.8  0.6  4.8 79.2 78.3  0.5  4.2 83.0
2017       74.4  0.6  4.6 79.6 78.8  0.5  4.1 83.3
```

To compensate the effect of the increasing life expectancy, we also compute the expected *percentage* of life that is spent with T1D, respectively T2D:

```
> Pj <- sweep( Sj, 1:2, apply( Sj, 1:2, sum ), "/" )
> Pj[,,1] <- Pj[,,2] + Pj[,,3]
> Pj <- Pj[,,c(2,3,1)]
> dimnames(Pj)[[3]][3] <- "DM"
> str( Pj )
 num [1:2, 1:22, 1:3] 0.00762 0.00624 0.00729 0.00599 0.007 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex  : chr [1:2] "M" "W"
  ..$ date : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ state: chr [1:3] "T1" "T2" "DM"
> round( ftable( Pj*100, row.vars=2 ), 1 )
     sex     M           W
     state  T1  T2  DM  T1  T2  DM
date
1996       0.8 2.8 3.6 0.6 2.6 3.2
1997       0.7 3.0 3.7 0.6 2.7 3.3
1998       0.7 3.1 3.8 0.6 2.8 3.4
1999       0.7 3.2 3.9 0.6 2.9 3.5
2000       0.7 3.4 4.0 0.5 3.0 3.6
2001       0.6 3.5 4.2 0.5 3.2 3.7
2002       0.6 3.7 4.3 0.5 3.3 3.8
2003       0.6 3.9 4.5 0.5 3.5 4.0
2004       0.6 4.1 4.8 0.5 3.7 4.2
2005       0.7 4.4 5.1 0.5 4.0 4.5
2006       0.7 4.8 5.5 0.5 4.3 4.8
2007       0.7 5.2 5.9 0.5 4.6 5.2
2008       0.7 5.7 6.4 0.6 5.0 5.5
```

```
2009       0.7 6.2 6.9 0.6 5.3 5.9
2010       0.7 6.5 7.2 0.6 5.6 6.2
2011       0.7 6.7 7.5 0.6 5.7 6.3
2012       0.7 6.8 7.5 0.6 5.7 6.3
2013       0.7 6.7 7.4 0.6 5.6 6.2
2014       0.7 6.5 7.2 0.6 5.4 6.0
2015       0.7 6.2 7.0 0.6 5.3 5.9
2016       0.7 6.0 6.7 0.6 5.1 5.7
2017       0.7 5.8 6.5 0.6 4.9 5.5
```

In order to create a table for the article we load the table of year of life lost:

```
> LR <- cPr[,paste(seq(1996,2017,3)),"100",1:2]
> str( LR )
 num [1:2, 1:8, 1:2] 1.98 1.53 1.72 1.3 1.56 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex  : chr [1:2] "M" "W"
  ..$ date : chr [1:8] "1996" "1999" "2002" "2005" ...
  ..$ state: chr [1:2] "T1" "T2"
> load( "../nydata/Yll.Rda" )
> str( YLL )
 num [1:2, 1:22, 1:2, 1:100, 1:5] 3.37 6.04 3.69 6.08 4 ...
 - attr(*, "dimnames")=List of 5
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ S.in: chr [1:2] "T1D" "T2D"
  ..$ A.in: chr [1:100] "0" "1" "2" "3" ...
  ..$ CoD : chr [1:5] "D-CVD" "D-Can" "D-Res" "D-Oth" ...
> yl1 <- YLL[,paste(seq(1996,2017,3)),"T1D",paste(1:3*20),"Sum"]
> yl2 <- YLL[,paste(seq(1996,2017,3)),"T2D",paste(2:4*20),"Sum"]
> str( yl1 )
 num [1:2, 1:8, 1:3] 9.74 11.21 11.6 11.83 12.93 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:8] "1996" "1999" "2002" "2005" ...
  ..$ A.in: chr [1:3] "20" "40" "60"
> str( yl2 )
 num [1:2, 1:8, 1:3] 9.46 9.54 8.71 8.65 8.05 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex : chr [1:2] "M" "W"
  ..$ date: chr [1:8] "1996" "1999" "2002" "2005" ...
  ..$ A.in: chr [1:3] "40" "60" "80"
> str( Sj )
 num [1:2, 1:22, 1:3] 70.2 75.5 70.4 75.7 70.7 ...
 - attr(*, "dimnames")=List of 3
  ..$ sex  : chr [1:2] "M" "W"
  ..$ date : chr [1:22] "1996" "1997" "1998" "1999" ...
  ..$ state: chr [1:3] "noDM" "T1" "T2"
> Sjx <- Sj[,paste(seq(1996,2017,3)),c(1:2,rep(1:3,3))]
> dimnames(Sjx)[[3]][1:9] <- c("LR:T1","LR:T2",
+                              paste("T1:",dimnames(yl1)[[3]],sep=""),
+                              paste("T2:",dimnames(yl2)[[3]],sep=""),
+                              "Sj: noDM")
> dimnames( Sjx )
```

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

$date
[1] "1996" "1999" "2002" "2005" "2008" "2011" "2014" "2017"

$state
 [1] "LR:T1"    "LR:T2"    "T1:20"    "T1:40"    "T1:60"    "T2:40"    "T2:60"
 [8] "T2:80"    "Sj: noDM" "T1"       "T2"
```

```
> Sjx[,,1:2  ] <- LR
> Sjx[,,1:3+2] <- yl1
> Sjx[,,1:3+5] <- yl2
> names( dimnames( Sjx ) )[3] <- ""
> round( ftable( Sjx, row.vars=1:2 ), 1 )
```

```
          LR:T1 LR:T2 T1:20 T1:40 T1:60 T2:40 T2:60 T2:80 Sj: noDM   T1   T2
sex date
M   1996    2.0  17.5   9.7   8.3   4.7   9.5   5.7   2.0     70.2  0.6  2.1
    1999    1.7  19.0  11.6  10.0   6.1   8.7   5.2   1.9     70.9  0.5  2.4
    2002    1.6  20.6  12.9  11.1   6.9   8.1   4.9   1.8     71.5  0.5  2.8
    2005    1.5  23.3  13.5  11.6   7.3   7.3   4.4   1.7     71.8  0.5  3.4
    2008    1.5  28.1  13.9  12.0   7.8   6.4   3.9   1.4     71.4  0.5  4.4
    2011    1.4  31.5  12.6  11.0   7.5   5.7   3.5   1.3     71.4  0.6  5.2
    2014    1.3  29.7  10.4   9.1   6.4   5.6   3.6   1.4     72.7  0.6  5.1
    2017    1.2  26.4   8.2   7.3   5.2   5.7   3.7   1.6     74.4  0.6  4.6
W   1996    1.5  15.5  11.2   9.7   6.4   9.5   6.3   2.5     75.5  0.5  2.0
    1999    1.3  16.5  11.8  10.3   7.1   8.6   5.7   2.3     75.9  0.4  2.3
    2002    1.1  17.7  12.2  10.8   7.7   7.8   5.1   2.1     76.3  0.4  2.6
    2005    1.1  19.8  12.7  11.2   8.1   7.0   4.6   1.9     76.5  0.4  3.2
    2008    1.1  23.4  12.9  11.4   8.4   6.2   4.0   1.6     76.2  0.4  4.0
    2011    1.0  25.7  11.7  10.4   7.9   5.7   3.7   1.5     76.3  0.5  4.7
    2014    1.0  23.9  10.0   9.0   6.9   5.6   3.7   1.5     77.4  0.5  4.5
    2017    1.0  21.1   8.4   7.6   6.0   5.7   3.8   1.6     78.8  0.5  4.1
```

```
> round( ftable( Sjx, row.vars=1:2 ), 2 )
```

```
          LR:T1 LR:T2 T1:20 T1:40 T1:60 T2:40 T2:60 T2:80 Sj: noDM    T1    T2
sex date
M   1996   1.98 17.53  9.74  8.32  4.73  9.46  5.71  2.04    70.22  0.56  2.07
    1999   1.72 18.96 11.60  9.99  6.06  8.71  5.25  1.89    70.87  0.50  2.39
    2002   1.56 20.62 12.93 11.14  6.95  8.05  4.86  1.78    71.47  0.48  2.77
    2005   1.51 23.32 13.52 11.61  7.33  7.31  4.45  1.65    71.77  0.50  3.36
    2008   1.50 28.05 13.86 11.95  7.85  6.36  3.89  1.41    71.44  0.53  4.36
    2011   1.43 31.54 12.61 10.97  7.51  5.72  3.53  1.30    71.43  0.55  5.21
    2014   1.32 29.71 10.40  9.13  6.42  5.64  3.55  1.41    72.73  0.57  5.06
    2017   1.25 26.42  8.24  7.28  5.18  5.72  3.68  1.58    74.38  0.59  4.62
W   1996   1.53 15.53 11.21  9.68  6.35  9.54  6.35  2.53    75.54  0.49  2.04
    1999   1.30 16.48 11.83 10.34  7.13  8.65  5.71  2.31    75.90  0.44  2.30
    2002   1.15 17.67 12.21 10.75  7.67  7.83  5.13  2.10    76.25  0.42  2.64
    2005   1.09 19.77 12.66 11.16  8.13  7.00  4.55  1.86    76.47  0.42  3.17
    2008   1.08 23.37 12.90 11.38  8.43  6.18  3.99  1.60    76.24  0.45  4.02
    2011   1.04 25.67 11.70 10.40  7.88  5.70  3.70  1.47    76.29  0.47  4.66
    2014   1.00 23.87  9.99  8.96  6.95  5.63  3.72  1.50    77.38  0.49  4.49
    2017   0.97 21.11  8.37  7.58  5.99  5.68  3.82  1.56    78.75  0.51  4.09
```

```
> a  <- fCtable( Sjx, row.vars=1:2, d=1, w=5 )
> a2 <- fCtable( Sjx, row.vars=1:2, d=2, w=5 )
> a[,c(1,10)] <- a2[,c(1,10)]
> ftable(a)
```

		LR:T1	LR:T2	T1:20	T1:40	T1:60	T2:40	T2:60	T2:80	Sj: noDM	T1	T2
sex	date											
M	1996	1.98	17.5	9.7	8.3	4.7	9.5	5.7	2.0	70.2	0.56	2.1
	1999	1.72	19.0	11.6	10.0	6.1	8.7	5.2	1.9	70.9	0.50	2.4
	2002	1.56	20.6	12.9	11.1	6.9	8.1	4.9	1.8	71.5	0.48	2.8
	2005	1.51	23.3	13.5	11.6	7.3	7.3	4.4	1.7	71.8	0.50	3.4
	2008	1.50	28.1	13.9	12.0	7.8	6.4	3.9	1.4	71.4	0.53	4.4
	2011	1.43	31.5	12.6	11.0	7.5	5.7	3.5	1.3	71.4	0.55	5.2
	2014	1.32	29.7	10.4	9.1	6.4	5.6	3.6	1.4	72.7	0.57	5.1
	2017	1.25	26.4	8.2	7.3	5.2	5.7	3.7	1.6	74.4	0.59	4.6
W	1996	1.53	15.5	11.2	9.7	6.4	9.5	6.3	2.5	75.5	0.49	2.0
	1999	1.30	16.5	11.8	10.3	7.1	8.6	5.7	2.3	75.9	0.44	2.3
	2002	1.15	17.7	12.2	10.8	7.7	7.8	5.1	2.1	76.3	0.42	2.6
	2005	1.09	19.8	12.7	11.2	8.1	7.0	4.6	1.9	76.5	0.42	3.2
	2008	1.08	23.4	12.9	11.4	8.4	6.2	4.0	1.6	76.2	0.45	4.0
	2011	1.04	25.7	11.7	10.4	7.9	5.7	3.7	1.5	76.3	0.47	4.7
	2014	1.00	23.9	10.0	9.0	6.9	5.6	3.7	1.5	77.4	0.49	4.5
	2017	0.97	21.1	8.4	7.6	6.0	5.7	3.8	1.6	78.8	0.51	4.1

```
----------------------
2019-03-27 at 16:50:49
Time elapsed: 00:00:35
----------------------
```

References

[1] B Carstensen. Age-Period-Cohort models for the Lexis diagram. *Statistics in Medicine*, 26(15):3018–3045, July 2007.

[2] TR Holford. The estimation of age, period and cohort effects for vital rates. *Biometrics*, 39:311–324, 1983.

[3] B Møller, H Fekjær, T Hakulinen, H Sigvaldason, HH Storm, M Talbäck, and T Haldorsen. Prediction of cancer incidence in the Nordic contries: empirical comparison of different approaches. *Statistics in Medicine*, 22:2751–2766, 2003.

[4] M. J. Rutherford, J. R. Thompson, and P. C. Lambert. Projecting cancer incidence using age-period-cohort models incorporating restricted cubic splines. *Int J Biostat*, 8(1):33, Nov 2012.

[5] Peter Sasieni. Age-period-cohort models in Stata. *Stata Journal*, 12:46–60, 2012.