Steno 2 study
20 year follow-up — end of 2014

SDC
March 2016

Version 3.6

Compiled Monday 6th June, 2016, 18:31
from: `/home/bendix/sdc/coll/jchq/r/steno2.tex`

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

Contents

Chapter 1

Data

1.1 Clinical visits

This file has data on the clinical visits that the trial participants have attended.

First we read the data and strip the irrelevant attributes:

```
> library( Epi )
> library( foreign )
> library(readstata13)
> clear()
> mic  <- read.dta13( "../data/micro2.dta", nonint.factors=TRUE )
```

We now construct a groomed base analysis data set for the clinical visits — mainly the microvascular outcomes:

```
> ( na <- names( attributes(mic) ) )
 [1] "row.names"        "names"            "datalabel"        "time.stamp"
 [5] "formats"          "types"            "val.labels"       "var.labels"
 [9] "version"          "label.table"      "expansion.fields" "strl"
[13] "byteorder"        "class"
> wh <- match( c("names","row.names","class","var.labels"), na )
> attributes( mic )[-wh] <- NULL
> dn <- grep( "date", names(mic) )
> dn <- dn[c(1,4,5,6,10,11,13,15)]
> # A few things that need to be corrected
> str( mic[,dn] )
'data.frame':        748 obs. of  8 variables:
 $ date          : Date, format: "1993-05-11" "1995-02-28" ...
 $ date_birth    : Date, format: "1932-04-20" "1932-04-20" ...
 $ date_blind    : Date, format: NA NA ...
 $ date_death    : Date, format: NA NA ...
 $ date_end      : Date, format: "2014-10-12" "2014-10-12" ...
 $ date_esrd     : Date, format: NA NA ...
 $ date_photocoag: num  NA NA NA NA NA ...
 $ dateeye2014   : Date, format: "2014-10-21" "2014-10-21" ...
> levels( mic$eds_r)
[1] "No retinopathy"            "Minimal non-proliferative" "Moderate nonproliferative"
[4] "Pre-proliferative"         "Photocoagulation"          "Proliferative"
> levels( mic$eds_l)
[1] "No retinopathy"            "Minimal non-proliferative" "Moderate nonproliferative"
[4] "Pre-proliferative"         "Photocoagulation"          "Proliferative"
> # transform.data.frame purges attributes so save them and put them back
> mat <- attributes(mic)
> mic <- transform( mic, date_photocoag = as.Date(date_photocoag,origin="1960-01-01"),
+                        eds_r = Relevel( eds_r, c(1:4,6:5) ),
+                        eds_l = Relevel( eds_l, c(1:4,6:5) ) )
> mat -> attributes(mic)
> names( mic )[dn]
```

```
[1] "date"           "date_birth"     "date_blind"     "date_death"     "date_end"
[6] "date_esrd"      "date_photocoag" "dateeye2014"
```

```
> nnam <- c("doVis",
+           "doBth",
+           "doBlind",
+           "doDth",
+           "doEnd",
+           "doESRD",
+           "doPhC",
+           "doEye")
> cbind( names( mic )[dn],   nnam )
```

```
                     nnam
[1,] "date"           "doVis"
[2,] "date_birth"     "doBth"
[3,] "date_blind"     "doBlind"
[4,] "date_death"     "doDth"
[5,] "date_end"       "doEnd"
[6,] "date_esrd"      "doESRD"
[7,] "date_photocoag" "doPhC"
[8,] "dateeye2014"    "doEye"
```

```
>        names( mic )[dn] <- nnam
> cbind( names(mic),
+   attributes(mic)$var.labels )
```

```
      [,1]              [,2]
 [1,] "ID"              "Patient ID"
 [2,] "SOPNR"           "SOPNR"
 [3,] "_merge"          ""
 [4,] "aer_level"       "u-AER class"
 [5,] "allocation"      "Treatment allocation"
 [6,] "autoprog"        "Progression in autonomic neuropathy"
 [7,] "blind"           "blindness according to WHO criteria"
 [8,] "doVis"           "Visit date"
 [9,] "date_aerprog"    ""
[10,] "date_autoprog"   ""
[11,] "doBth"           "Date of birth"
[12,] "doBlind"         "Date of first eye-examination with blindness"
[13,] "doDth"           "Date of death"
[14,] "date_eds_l_prog" ""
[15,] "date_eds_prog"   ""
[16,] "date_eds_r_prog" ""
[17,] "doEnd"           "Date of last observation"
[18,] "doESRD"          "Date of Renal Replacement Therapy initiation"
[19,] "date_periprog"   ""
[20,] "doPhC"           ""
[21,] "date_progeds"    ""
[22,] "doEye"           "Date of latest retinography after 2006"
[23,] "eds"             ""
[24,] "eds_l"           "EURO-DIAB score, left"
[25,] "eds_l_prog"      ""
[26,] "eds_max"         ""
[27,] "eds_r"           "EURO-DIAB score, right"
[28,] "eds_r_prog"      ""
[29,] "esrd"            "End Stage Renal Disease"
[30,] "if_aerprog"      ""
[31,] "if_autoprog"     ""
[32,] "if_periprog"     ""
[33,] "if_photocoag"    ""
[34,] "laser"           "Photocoagulation"
[35,] "mean_aer"        "Geometric mean of u-AER"
[36,] "med_aer"         "Median u-AER"
[37,] "periprog"        "Progression in peripheral neuropathy"
[38,] "prog_eds_max"    ""
[39,] "sex"             "Gender"
[40,] "year"            "Follow-up point"
```

```
> mic <- cal.yr( mic )
> str( mic )
'data.frame':        748 obs. of  40 variables:
 $ ID              : num  1 1 1 1 1 1 2 2 2 2 ...
 $ SOPNR           : num  1 1 1 1 1 1 2 2 2 2 ...
 $ _merge          : Factor w/ 5 levels "master only (1)",..: 3 3 3 3 3 3 3 3 3 3 ...
 $ aer_level       : Factor w/ 3 levels "Normoalbuminuria",..: 2 2 1 2 1 2 2 1 2 3 ...
 $ allocation      : Factor w/ 2 levels "Intensive","Conventional": 1 1 1 1 1 1 1 1 1 1 ...
 $ autoprog        : int  NA 0 0 0 0 0 NA 1 1 1 ...
 $ blind           : int  0 NA 0 0 0 0 0 NA 0 0 ...
 $ doVis           :Classes 'cal.yr', 'numeric'  num [1:748] 1993 1995 1997 2001 2007 ...
 $ date_aerprog    : num  NA NA NA NA NA ...
 $ date_autoprog   : num  NA NA NA NA NA ...
 $ doBth           :Classes 'cal.yr', 'numeric'  num [1:748] 1932 1932 1932 1932 1932 ...
 $ doBlind         :Classes 'cal.yr', 'numeric'  num [1:748] NA NA NA NA NA NA NA NA NA NA ...
 $ doDth           :Classes 'cal.yr', 'numeric'  num [1:748] NA NA NA NA NA NA NA NA NA NA ...
 $ date_eds_l_prog: num  NA NA NA NA NA ...
 $ date_eds_prog   : num  NA NA NA NA NA ...
 $ date_eds_r_prog: num  NA NA NA NA NA ...
 $ doEnd           :Classes 'cal.yr', 'numeric'  num [1:748] 2015 2015 2015 2015 2015 ...
 $ doESRD          :Classes 'cal.yr', 'numeric'  num [1:748] NA NA NA NA NA NA NA NA NA NA ...
 $ date_periprog   : num  NA NA NA NA NA ...
 $ doPhC           :Classes 'cal.yr', 'numeric'  num [1:748] NA NA NA NA NA ...
 $ date_progeds    : num  NA NA NA NA NA ...
 $ doEye           :Classes 'cal.yr', 'numeric'  num [1:748] 2015 2015 2015 2015 2015 ...
 $ eds             : num  NA NA NA NA NA 1 NA NA NA NA ...
 $ eds_l           : Factor w/ 6 levels "No retinopathy",..: 1 NA 1 1 1 NA 2 NA 2 6 ...
 $ eds_l_prog      : int  NA NA 0 0 0 NA NA NA 0 1 ...
 $ eds_max         : num  0 NA 0 0 0 1 1 NA 1 4 ...
 $ eds_r           : Factor w/ 6 levels "No retinopathy",..: 1 NA 1 1 1 NA 2 NA 2 3 ...
 $ eds_r_prog      : int  NA NA 0 0 0 NA NA NA 0 1 ...
 $ esrd            : int  0 0 0 0 0 0 0 0 0 0 ...
 $ if_aerprog      : int  0 0 0 0 0 0 0 0 0 1 ...
 $ if_autoprog     : int  0 0 0 0 0 0 0 1 1 1 ...
 $ if_periprog     : int  0 0 0 0 0 0 0 0 0 1 ...
 $ if_photocoag    : int  0 0 0 0 0 0 0 0 0 1 ...
 $ laser           : int  NA NA NA 0 0 NA NA NA NA 1 ...
 $ mean_aer        : num  76.9 34.9 21 92.8 19.5 ...
 $ med_aer         : num  76.5 34 23 105 18 83 69.5 17 86 328 ...
 $ periprog        : int  NA NA 0 0 0 0 NA NA 0 1 ...
 $ prog_eds_max    : int  0 NA 0 0 0 1 0 NA 0 1 ...
 $ sex             : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 2 2 2 2 ...
 $ year            : int  1993 1995 1997 2001 2006 2014 1993 1995 1997 2001 ...
 - attr(*, "var.labels")= chr  "Patient ID" "SOPNR" "" "u-AER class" ...
> # How many have how many visits
> addmargins( t( apply( table(mic$ID,mic$allocation), 2, table ) )[,-1] )
              1 2  3  4  5  6 Sum
  Intensive   2 2 10 12 15 39  80
  Conventional 1 3 13 27 15 21  80
  Sum         3 5 23 39 30 60 160
```

We devise a variable `visit` that just numbers the visits from 1 to 6, and also plot the visit dates versus the visit numbers:

```
> mic$visit <- Relevel( factor(floor(mic$doVis)),
+                       list("1"=1:2,
+                            "2"=3,
+                            "3"=4,
+                            "4"=5,
+                            "5"=6:7,
+                            "6"=8:9) )
> mic$visit <- as.numeric( as.character( mic$visit ) )
> # Check that we actually got the numbering correct:
> print( with( mic, addmargins( table( floor(doVis*5)/5, visit, useNA="ifany" ),
+                               margin=1 ) ),
+        zero.print="." )
```

```
        visit
          1   2   3   4   5   6
  1993   35   .   .   .   .   .
  1993.2 43   .   .   .   .   .
  1993.4 16   .   .   .   .   .
  1993.6 19   .   .   .   .   .
  1993.8 16   .   .   .   .   .
  1994   21   .   .   .   .   .
  1994.2 10   .   .   .   .   .
  1995    .  50   .   .   .   .
  1995.2  .  40   .   .   .   .
  1995.4  .   8   .   .   .   .
  1995.6  .  46   .   .   .   .
  1995.8  .  12   .   .   .   .
  1997    .   .  52   .   .   .
  1997.2  .   .  33   .   .   .
  1997.4  .   .  10   .   .   .
  1997.6  .   .  45   .   .   .
  1997.8  .   .   9   .   .   .
  2001    .   .   .  33   .   .
  2001.2  .   .   .  38   .   .
  2001.4  .   .   .   1   .   .
  2001.6  .   .   .  39   .   .
  2001.8  .   .   .  17   .   .
  2006.6  .   .   .   .  38   .
  2006.8  .   .   .   .  47   .
  2007    .   .   .   .   8   .
  2014    .   .   .   .   .   2
  2014.2  .   .   .   .   .   2
  2014.4  .   .   .   .   .   9
  2014.6  .   .   .   .   .  30
  2014.8  .   .   .   .   .  17
  2015    .   .   .   .   .   2
  Sum   160 156 149 128  93  62
```

```
> with( mic, plot( doVis,
+                  visit+ID/300-88/300,
+                  pch=16, cex=0.5,
+                  xlab="Date of examination",
+                  ylab="Visit no" ) )
> abline( v=1990:2015, h=seq(-0.5,5.5,1), col=gray(0.8) )
```

Finally we restrict the visits dataset to the relevant variables:

```
> mic <- mic[,c("ID","allocation","sex",
+               "doBth","doVis","doESRD","doEye","doPhC","doBlind","doEnd","doDth",
+               "visit","aer_level","med_aer","eds_l","eds_r","autoprog","periprog")]
> save( mic, file="../data/mic.Rda" )
```

For the sake of analysis of ESRD occurrence we make a small dataset with ID and ESRD dates for merging with the baseline data set:

```
> zz <- with( mic, tapply( doESRD, ID, mean, na.rm=TRUE ) )
> ESRD <- data.frame( ID = as.numeric(names(zz)),
+                     doESRD = as.numeric(zz) )
> str(ESRD)
'data.frame':        160 obs. of  2 variables:
 $ ID    : num  1 2 3 4 5 6 7 8 9 10 ...
 $ doESRD: num  NaN NaN NaN NaN 1998 ...

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

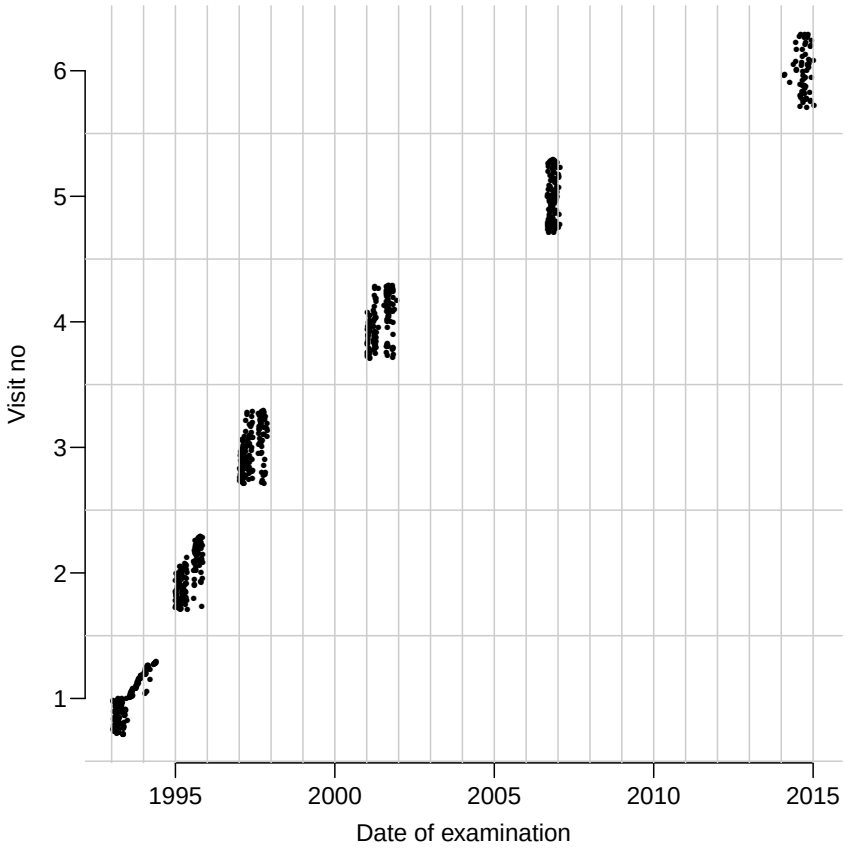

Figure 1.1: *Visit number (plus ID/200) versus date of visit; thus the persons are vertically ordered by ID within each visit.*

1.2 Event data

We read the supplied event data, groom the variable names, assign a date of DM diagnosis in the interval supplied — thus we assign a random date between 0 and 1 years prior to baseline date for those coded 0 on `dmvar93`, a random date between 1 and 2 years prior to baseline date for those coded 1 on `dmvar93`, etc. :

```
> library( foreign )
> library( Epi )
> clear()
> st2 <- read.dta( "../data/BXC til multistate stata12.dta" )
> ### Change funny characters in variable names
> names(st2)[c(14,16)]
[1] "ev1_inkl_d\xf8d"     "event_inkl_CVd\xf8d"
> names(st2)[c(14,16)] <- c("ev1_inkl_dod","event_inkl_CVdod")
> ## Convert date to class cal.yr
> st2 <- cal.yr( st2 )
> ## Merge with updated info on baseline cvd
> load( "../data/bcvd.Rda" )
> st2 <- merge( st2, bcvd )
> ## Merge with ESRD date
> load( "../data/ESRD.Rda" )
> st2 <- merge( st2, ESRD )
> st2 <- transform( st2, allocation = factor( allocation,
+                                             labels=c("Intensive","Conventional") ),
+                                sex = factor( sex,
+                                             labels=c("F","M") ) )
```

```
> ## Change name of date variables
> names(st2) <- gsub("date.","do",gsub("_",".",names(st2)))
> # Assign a date of DM diagnosis, randomly in the year given
> set.seed(78468)
> st2$dodm <- st2$dobaseline - ( st2$dmvar93 + runif(nrow(st2)) )
```

Now rename the date variables to something a bit more mnemonic:

```
> names(st2)[wh <- substr(names(st2),1,2)=="do"]
[1] "dobaseline" "dobirth"    "dodeath"    "doend"      "doevent1"   "doevent2"
[7] "doevent3"   "doESRD"     "dodm"
> newnam <- c("doBase","doBth","doDth","doEnd","doCVD1","doCVD2","doCVD3","doESRD","doDM")
> cbind( names(st2)[wh],   newnam )
                 newnam
 [1,] "dobaseline" "doBase"
 [2,] "dobirth"    "doBth"
 [3,] "dodeath"    "doDth"
 [4,] "doend"      "doEnd"
 [5,] "doevent1"   "doCVD1"
 [6,] "doevent2"   "doCVD2"
 [7,] "doevent3"   "doCVD3"
 [8,] "doESRD"     "doESRD"
 [9,] "dodm"       "doDM"
>        names(st2)[wh] <- newnam
> names(st2)[wh <- grep("cvd",names(st2))]
[1] "base.cvd"  "cvd.death"
> newnam <- c("baseCVD","deathCVD")
> cbind( names(st2)[wh],   newnam )
               newnam
[1,] "base.cvd"  "baseCVD"
[2,] "cvd.death" "deathCVD"
>        names(st2)[wh] <- newnam
```

Finally we subset and reorder the variables — note that we for conveniece have two versions of the allocation variable with short and long levels names.

```
> st2 <- st2[,c("ID","allocation","sex","baseCVD","deathCVD",
+               "doBth","doDM","doBase","doCVD1","doCVD2","doCVD3","doESRD","doEnd","doDth")]
```

1.2.1 Checking the date variables

We make pairwise plot of all date variables, in the logical order of the variables so that we can see that all points should be on the same side of the red identity lines in figure 1.2:

```
> names(st2)[dvar <- grep("do",names(st2))]
[1] "doBth"  "doDM"   "doBase" "doCVD1" "doCVD2" "doCVD3" "doESRD" "doEnd"  "doDth"
> par( bty="o" )
> pairs( st2[,dvar], gap=0,
+        panel=function(x,y){points(x,y,pch=16,cex=0.8)
+                            abline(0,1,col="red")} )
> par( bty="n" )
```

From the pairwise plots it is seen that the end-dates do not have a single value at the end where all persons alive are censored; the date in the variable `doEnd` is the date of last examination.

It seems as if some of the event dates for successive events are identical, but this is actually not the case; here is a list of all (CVD)event dates recorded after baseline, and of the differences (in years) between successive CVD event dates:

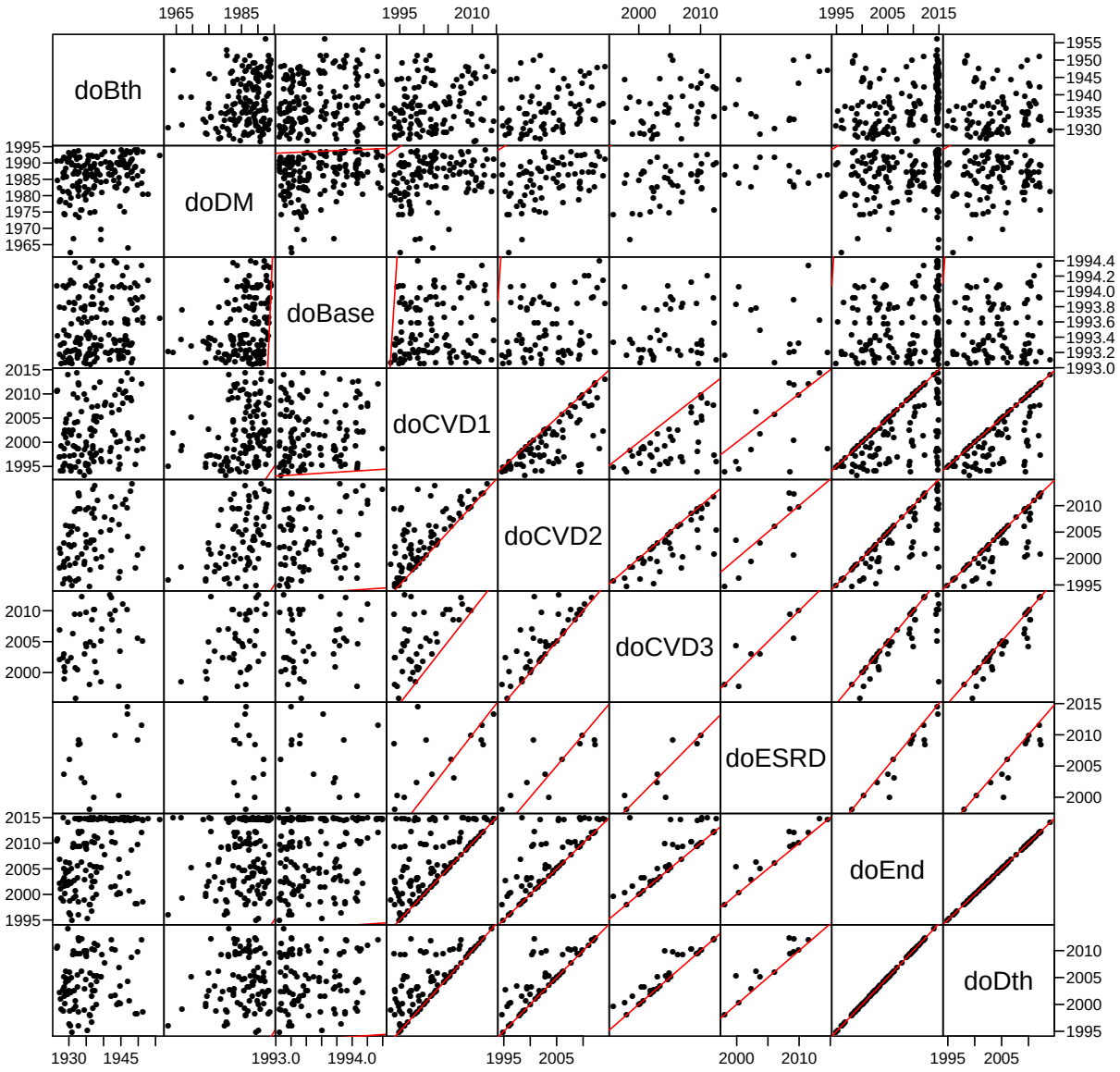

Figure 1.2: *Pairwise plots of all dates in the data frame. The red lines are the identity lines. With the exception of* `doCVD` *and* `doESRD`*, points in the upper panels should be right / below the identity lines.*

```
> ev <- st2[,c("doBase","doCVD1","doCVD2","doCVD3","doEnd","doDth")]
> head(ev)
    doBase   doCVD1   doCVD2   doCVD3    doEnd    doDth
1 1993.357 2014.457       NA       NA 2014.778       NA
2 1993.357 2009.231 2009.570 2010.162 2014.589       NA
3 1993.362 2001.639       NA       NA 2001.639 2001.639
4 1993.362 1995.415 1997.461 2003.443 2003.443 2003.443
5 1993.162 1993.959 1994.657 1998.047 1998.047 1998.047
6 1993.201 1998.769       NA       NA 2014.684       NA
> wh <- apply(ev[,2:4],1,function(x)sum(is.na(x)))<3
> round( cbind( ev[wh,], t(apply(ev[wh,2:4],1,diff)) ), 3 )
     doBase   doCVD1   doCVD2   doCVD3    doEnd    doDth doCVD2 doCVD3
1  1993.357 2014.457       NA       NA 2014.778       NA     NA     NA
```

2	1993.357	2009.231	2009.570	2010.162	2014.589	NA	0.339	0.591
3	1993.362	2001.639	NA	NA	2001.639	2001.639	NA	NA
4	1993.362	1995.415	1997.461	2003.443	2003.443	2003.443	2.045	5.982
5	1993.162	1993.959	1994.657	1998.047	1998.047	1998.047	0.698	3.389
6	1993.201	1998.769	NA	NA	2014.684	NA	NA	NA
7	1993.203	2001.058	2006.238	NA	2006.238	2006.238	5.180	NA
8	1993.209	2011.925	2012.149	NA	2012.149	2012.149	0.225	NA
9	1993.209	1995.123	1995.952	NA	1995.952	1995.952	0.830	NA
10	1993.075	2011.405	NA	NA	2011.405	2011.405	NA	NA
11	1993.066	2005.666	2006.068	NA	2006.068	2006.068	0.402	NA
12	1993.146	1994.000	1995.306	2002.427	2002.739	2002.739	1.306	7.121
14	1993.105	2014.038	NA	NA	2014.038	2014.038	NA	NA
16	1993.053	2000.034	2002.112	2009.680	2009.680	2009.680	2.078	7.567
18	1993.373	1998.260	1998.471	1998.594	2014.997	NA	0.211	0.123
21	1993.379	1994.274	NA	NA	2014.789	NA	NA	NA
23	1993.069	1993.132	1998.994	NA	1998.994	1998.994	5.862	NA
24	1993.187	2005.431	2007.667	2010.156	2010.156	2010.156	2.237	2.489
26	1993.225	1999.780	1999.985	2001.855	2001.855	2001.855	0.205	1.870
27	1993.222	2002.531	2006.175	2006.334	2009.228	2009.228	3.644	0.159
29	1993.280	2005.222	NA	NA	2005.222	2005.222	NA	NA
30	1993.392	2004.300	2005.543	NA	2005.543	2005.543	1.243	NA
31	1993.220	2006.112	2007.235	2008.513	2010.342	2010.342	1.123	1.279
32	1993.220	2000.168	NA	NA	2000.168	2000.168	NA	NA
34	1993.105	1996.045	2005.431	2012.546	2014.745	NA	9.385	7.116
35	1993.496	2001.710	2002.865	2003.018	2005.091	2005.091	1.155	0.153
36	1993.203	2001.912	NA	NA	2014.901	NA	NA	NA
37	1993.069	2008.697	NA	NA	2008.697	2008.697	NA	NA
38	1993.129	1998.742	2000.198	2000.407	2003.322	2003.322	1.457	0.208
39	1993.110	2010.838	NA	NA	2014.654	NA	NA	NA
40	1993.239	1997.765	2000.012	2000.012	2000.012	2000.012	2.248	0.000
41	1993.066	2002.471	NA	NA	2002.471	2002.471	NA	NA
42	1993.244	2002.758	2002.928	2004.902	2005.838	2005.838	0.170	1.974
43	1993.107	1994.052	2002.181	2002.367	2002.367	2002.367	8.129	0.186
45	1993.132	2010.233	NA	NA	2014.745	NA	NA	NA
46	1993.398	1997.409	2001.715	2001.869	2002.517	2002.517	4.307	0.153
47	1993.414	2010.789	NA	NA	2010.789	2010.789	NA	NA
48	1993.417	2002.589	2002.605	NA	2002.605	2002.605	0.016	NA
49	1993.417	2010.564	2011.199	NA	2014.767	NA	0.635	NA
50	1993.168	1996.880	2006.402	2006.780	2014.665	NA	9.522	0.378
51	1993.181	1996.957	2001.904	2002.145	2002.145	2002.145	4.947	0.241
54	1993.206	2012.283	2012.305	NA	2012.305	2012.305	0.022	NA
55	1993.072	1995.602	1996.042	2000.223	2000.223	2000.223	0.441	4.181
56	1993.124	2005.387	NA	NA	2005.387	2005.387	NA	NA
57	1993.431	2004.719	NA	NA	2004.719	2004.719	NA	NA
58	1993.431	1997.647	1997.937	NA	1997.937	1997.937	0.290	NA
60	1993.433	1999.328	2003.035	2004.209	2009.975	2009.975	3.707	1.175
61	1993.242	1999.040	1999.139	2000.743	2000.743	2000.743	0.099	1.604
62	1993.143	2006.786	NA	NA	2006.786	2006.786	NA	NA
64	1993.170	1996.045	1996.045	NA	1996.045	1996.045	0.000	NA
65	1993.299	1994.093	2001.001	NA	2001.001	2001.001	6.908	NA
66	1993.296	2000.710	NA	NA	2000.710	2000.710	NA	NA
68	1993.113	1995.166	2009.203	NA	2009.203	2009.203	14.037	NA
69	1993.302	2004.626	2009.368	2010.181	2014.953	NA	4.742	0.813
71	1993.124	1997.080	1998.695	NA	1998.695	1998.695	1.615	NA
72	1993.318	1997.211	1998.400	1998.999	2001.595	2001.595	1.188	0.600
73	1993.151	2008.867	NA	NA	2008.867	2008.867	NA	NA
74	1993.181	1996.689	NA	NA	1996.689	1996.689	NA	NA
77	1993.335	1994.841	1995.684	1995.875	1999.577	1999.577	0.843	0.192
78	1993.206	2008.305	NA	NA	2014.997	NA	NA	NA
79	1993.231	1999.955	NA	NA	1999.955	1999.955	NA	NA
80	1993.050	1999.194	NA	NA	1999.194	1999.194	NA	NA
81	1993.050	1994.868	1994.873	NA	1994.873	1994.873	0.005	NA
82	1993.165	2009.233	2009.373	NA	2009.373	2009.373	0.140	NA
83	1993.321	2009.710	2009.841	2010.071	2010.071	2010.071	0.131	0.230
85	1993.316	1993.825	2009.422	2009.436	2009.444	2009.444	15.598	0.014
86	1993.201	1998.057	2005.318	2009.515	2014.479	NA	7.261	4.197

```
87  1993.469 1995.407 2009.636       NA 2014.805       NA 14.229     NA
88  1993.570 2001.521 2004.229       NA 2004.229 2004.229  2.708     NA
89  1993.570 2001.567 2004.998       NA 2004.998 2004.998  3.431     NA
91  1993.584 1998.621       NA       NA 1998.621 1998.621     NA     NA
92  1993.584 2007.728 2007.760       NA 2007.760 2007.760  0.033     NA
93  1993.661 2004.281       NA       NA 2004.281 2004.281     NA     NA
94  1993.587 1998.969 2001.499       NA 2001.499 2001.499  2.530     NA
96  1993.592 1997.340 2000.828 2012.223 2012.223 2012.223  3.488 11.395
99  1993.606 1995.243       NA       NA 1995.243 1995.243     NA     NA
101 1993.622 2014.457       NA       NA 2014.827       NA     NA     NA
103 1993.644 2001.450       NA       NA 2001.450 2001.450     NA     NA
105 1993.663 2009.890       NA       NA 2009.890 2009.890     NA     NA
108 1993.762 1999.188       NA       NA 1999.188 1999.188     NA     NA
109 1993.759 1998.496 1999.374 2002.964 2002.964 2002.964  0.879  3.589
111 1993.778 1998.608 1999.413 2000.872 2003.218 2003.218  0.805  1.459
112 1993.781 2006.260       NA       NA 2006.260 2006.260     NA     NA
113 1993.781 1998.230       NA       NA 1998.230 1998.230     NA     NA
114 1993.775 1996.505 2005.053 2005.179 2005.179 2005.179  8.548  0.126
115 1993.795 1997.001 1998.175 2006.895 2006.895       NA  1.175  8.720
116 1993.838 2013.097 2014.153       NA 2014.668       NA  1.057     NA
118 1993.836 1995.938 2003.552 2004.366 2005.381 2005.381  7.614  0.813
119 1993.836 2002.841 2003.522       NA 2003.522 2003.522  0.682     NA
120 1993.838 1996.064 2000.330 2007.095 2009.351 2009.351  4.266  6.765
121 1993.841 2003.988 2008.905 2009.376 2009.376 2009.376  4.917  0.471
122 1994.205 2007.963 2010.277 2011.109 2014.997       NA  2.313  0.832
123 1993.855 1994.975 1996.371       NA 1996.371 1996.371  1.396     NA
124 1993.855 1997.373 2002.632       NA 2002.632 2002.632  5.259     NA
125 1993.890 2000.483 2000.606 2005.529 2009.770 2009.770  0.123  4.923
126 1993.910 1998.632       NA       NA 1998.632 1998.632     NA     NA
128 1993.910 1995.503 1998.230       NA 1998.230 1998.230  2.727     NA
129 1993.929 2001.236 2001.921 2005.159 2014.665       NA  0.684  3.239
130 1993.915 2007.708 2011.643 2012.081 2012.081 2012.081  3.934  0.438
131 1993.932 2002.487       NA       NA 2002.487 2002.487     NA     NA
132 1993.929 2001.444 2011.914       NA 2011.914 2011.914 10.470     NA
133 1994.047 1999.240       NA       NA 1999.240 1999.240     NA     NA
134 1994.066 1994.496 1996.330 1997.723 2000.338 2000.338  1.834  1.394
135 1994.082 2012.713       NA       NA 2014.997       NA     NA     NA
136 1994.011 2012.522       NA       NA 2014.997       NA     NA     NA
138 1994.052 2004.122       NA       NA 2004.122 2004.122     NA     NA
139 1994.049 2009.214 2012.678       NA 2014.750       NA  3.463     NA
140 1994.074 1995.706 2004.442 2004.700 2004.700 2004.700  8.736  0.257
142 1994.066 2002.331       NA       NA 2014.457       NA     NA     NA
143 1994.205 2007.522 2011.076       NA 2011.076 2011.076  3.554     NA
144 1994.104 2009.967       NA       NA 2009.967 2009.967     NA     NA
145 1994.071 2010.465       NA       NA 2010.465 2010.465     NA     NA
146 1994.071 1994.983 1998.701       NA 1998.701 1998.701  3.718     NA
148 1994.085 2002.107 2004.439       NA 2004.439 2004.439  2.333     NA
149 1994.140 2002.186       NA       NA 2002.186 2002.186     NA     NA
150 1994.120 2007.314 2008.494 2008.494 2010.378 2010.378  1.180  0.000
152 1994.142 2002.266 2013.792       NA 2014.824       NA 11.526     NA
156 1994.339 2012.051       NA       NA 2012.051 2012.051     NA     NA
159 1994.397 1998.742 2013.099       NA 2014.611       NA 14.357     NA
```

1.3 Lexis objects of follow-up

In order to get an overview of follow-up of persons we set up a `Lexis` object with 4 timescales: calendar time, age, diabetes duration and time since baseline:

```
> names( st2 )
 [1] "ID"         "allocation" "sex"        "baseCVD"    "deathCVD"   "doBth"
 [7] "doDM"       "doBase"     "doCVD1"     "doCVD2"     "doCVD3"     "doESRD"
[13] "doEnd"      "doDth"
```

```
> L0 <- Lexis( entry = list( per = doBase,
+                            age = doBase-doBth,
+                            dur = doBase-doDM,
+                            tsb = 0 ),
+              exit = list( per = doEnd ),
+              exit.status = factor( (!is.na(doDth))+deathCVD,
+                                    labels=c("DM","Dead","CV-D") ),
+              data = st2 )

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

> summary( L0 )

Transitions:
     To
From DM Dead CV-D  Records:  Events: Risk time:  Persons:
  DM 67   55   38       160       93     2417.6       160
```

We then cut the follow-up at the CVD event dates:

```
> L1 <- cutLexis( L0, cut=L0$doCVD1,
+                     new.state="1st CVD", split.states=TRUE,
+                     precursor.states=c("DM") )
> L1 <- cutLexis( L1, cut=L1$doCVD2,
+                     new.state="2nd CVD", split.states=TRUE,
+                     precursor.states=c("DM","1st CVD") )
> L1 <- cutLexis( L1, cut=L1$doCVD3,
+                     new.state="3rd CVD", split.states=TRUE,
+                     precursor.states=c("DM","1st CVD","2nd CVD") )
```

When keeping track of the states *from* which death occurs, we get awfully long state names by default, so we groom these and finally reorder them to facilitate plotting and summarizing;

```
> levels( L1 )
 [1] "DM"                                "1st CVD"
 [3] "2nd CVD"                           "3rd CVD"
 [5] "Dead"                              "CV-D"
 [7] "Dead(1st CVD)"                     "CV-D(1st CVD)"
 [9] "Dead(1st CVD)(2nd CVD)"            "CV-D(1st CVD)(2nd CVD)"
[11] "CV-D(1st CVD)(2nd CVD)(3rd CVD)" "Dead(1st CVD)(2nd CVD)(3rd CVD)"

> # A function to tease out the number of previous events from the state
> # names, basically fishing out the 8th last character
> subrv8 <- function(x) ifelse( (wh<-(nchar(x)-7))>0, substr(x,wh,wh), "0" )
> cbind(
+ levels( L1$lex.Xst )[-(1:4)],
+ paste( substr( levels( L1$lex.Xst )[-(1:4)], 1, 4 ),
+        subrv8( levels( L1$lex.Xst )[-(1:4)] ) ) )
     [,1]                               [,2]
[1,] "Dead"                             "Dead 0"
[2,] "CV-D"                             "CV-D 0"
[3,] "Dead(1st CVD)"                    "Dead 1"
[4,] "CV-D(1st CVD)"                    "CV-D 1"
[5,] "Dead(1st CVD)(2nd CVD)"           "Dead 2"
[6,] "CV-D(1st CVD)(2nd CVD)"           "CV-D 2"
[7,] "CV-D(1st CVD)(2nd CVD)(3rd CVD)" "CV-D 3"
[8,] "Dead(1st CVD)(2nd CVD)(3rd CVD)" "Dead 3"

> levels( L1$lex.Cst )[-(1:4)] <-
+ levels( L1$lex.Xst )[-(1:4)] <-
+ paste( substr( levels( L1$lex.Xst )[-(1:4)], 1, 4 ),
+        subrv8( levels( L1$lex.Xst )[-(1:4)] ) )
> L1 <- Relevel( L1, c(1:10,12,11) )
> summary( L1 )
```

```
Transitions:
     To
From      DM 1st CVD 2nd CVD 3rd CVD Dead 0 CV-D 0 Dead 1 CV-D 1 Dead 2 CV-D 2 Dead 3
  DM      41      86       0       0     24      9      0      0      0      0      0
  1st CVD  0      11      48       0      0      0     14     13      0      0      0
  2nd CVD  0       0       7      24      0      0      0      0     12      4      0
  3rd CVD  0       0       0       8      0      0      0      0      0      0      5
  Sum     41      97      55      32     24      9     14     13     12      4      5

Transitions:
     To
From      CV-D 3  Records:  Events: Risk time:  Persons:
  DM           0       160      119    1870.64       160
  1st CVD      0        86       75     342.61        86
  2nd CVD      0        47       40     112.27        47
  3rd CVD     12        25       17      92.08        25
  Sum         12       318      251    2417.60       160
```

With this in order we can now show how persons fare through the study, separately for the two randomization groups:

```
> par( mfrow=c(1,2) )
> for( i in 1:2 ){
+ boxes( subset(L1,allocation==levels(allocation)[i]),
+            boxpos=list(x=c(rep(20,4),rep(80,8)),
+                        y=c(seq(90,10,,4),
+                        rep(seq(92,8,,4),each=2)+rep(c(1,-1)*4,4) ) ),
+            pos.arr=rep(c(5,4,6)/10,4)[-12],#[-15+i],
+            show.BE="nz", DR.sep=c(" (", ")"), eq.ht=FALSE, hmult=1.3,
+            scale.R=1000)
+ text(10,par("usr")[4],levels(L1$allocation)[i],adj=c(0,1),font=2,cex=2)
+ }
```

1.3.1 Follow-up after end of intervention

According to Gæde *et al.*[1] the trial intervention ended at a mean of 7.8 years after intervention, so we define this as the end of study:

```
> as.Date.cal.yr( doStEnd <- mean(st2$doBase) + 7.8 )
[1] "2001-05-09"
```

First we construct the dataset `P1` restricted to follow-up after end of study:

```
> S1 <- splitLexis( L1, breaks=doStEnd, time.scale="per" )
> S1[S1$lex.id<4,1:9]
   lex.id      per      age       dur       tsb    lex.dur lex.Cst lex.Xst ID
1       1 1993.357 61.05681  2.263567  0.000000  7.9942505      DM      DM  1
2       1 2001.351 69.05106 10.257817  7.994251 13.1063655      DM 1st CVD  1
3       1 2014.457 82.15743 23.364183 21.100616  0.3203285 1st CVD 1st CVD  1
4       2 1993.357 46.54346 11.226522  0.000000  7.9942505      DM      DM  2
5       2 2001.351 54.53771 19.220772  7.994251  7.8798084      DM 1st CVD  2
6       2 2009.231 62.41752 27.100581 15.874059  0.3394935 1st CVD 2nd CVD  2
7       2 2009.570 62.75702 27.440074 16.213552  0.5913758 2nd CVD 3rd CVD  2
8       2 2010.162 63.34839 28.031450 16.804928  4.4271047 3rd CVD 3rd CVD  2
9       3 1993.362 49.98494 10.801806  0.000000  7.9887748      DM      DM  3
10      3 2001.351 57.97372 18.790580  7.988775  0.2877481      DM  Dead 0  3
> P1 <- subset( S1, per>=doStEnd )
> P1[P1$lex.id<4,1:9]
   lex.id      per      age      dur       tsb    lex.dur lex.Cst lex.Xst ID
2       1 2001.351 69.05106 10.25782  7.994251 13.1063655      DM 1st CVD  1
3       1 2014.457 82.15743 23.36418 21.100616  0.3203285 1st CVD 1st CVD  1
5       2 2001.351 54.53771 19.22077  7.994251  7.8798084      DM 1st CVD  2
```

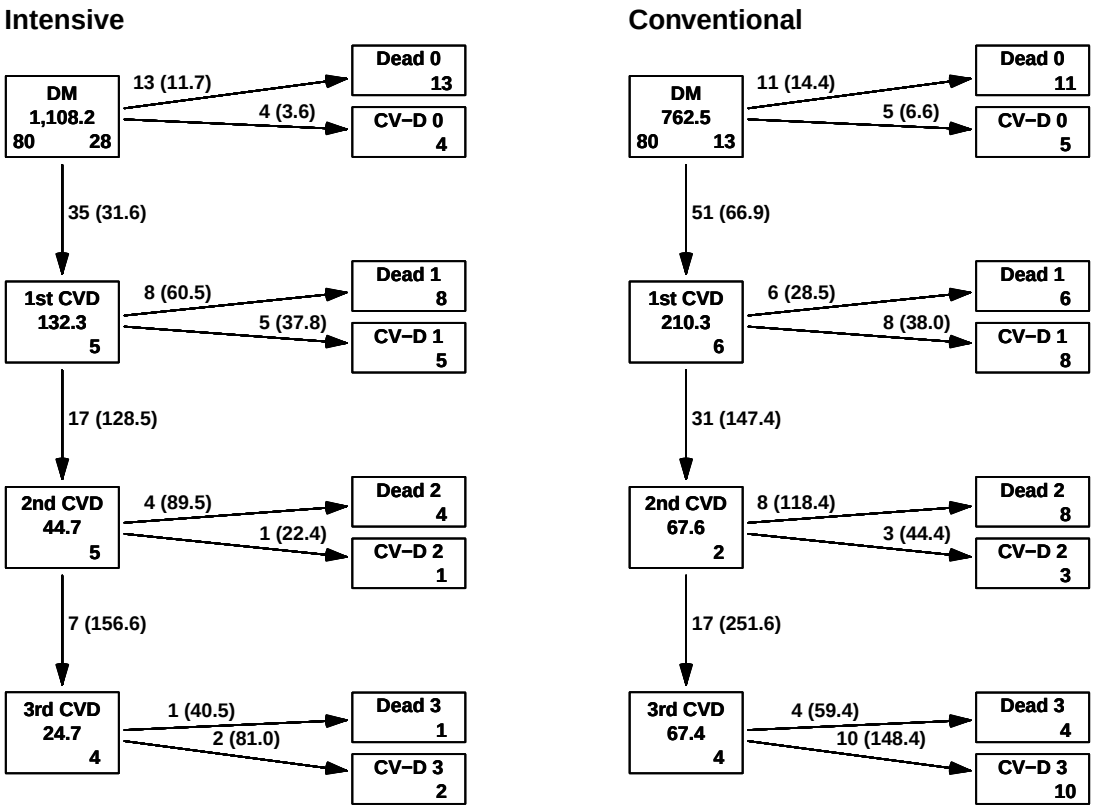

Figure 1.3: *Flow of patients from the Steno2 study through CVD states to death by intervention group.*
The numbers in the middle of the boxes are the person-years, the number at the bottom of the boxes are the number of persons starting (left), resp. ending (right) their follow-up in each state. The numbers on the arrows are the number of transitions (transition-rates per 1000 PY).

```
6       2 2009.231 62.41752 27.10058 15.874059  0.3394935 1st CVD 2nd CVD  2
7       2 2009.570 62.75702 27.44007 16.213552  0.5913758 2nd CVD 3rd CVD  2
8       2 2010.162 63.34839 28.03145 16.804928  4.4271047 3rd CVD 3rd CVD  2
10      3 2001.351 57.97372 18.79058  7.988775  0.2877481      DM  Dead 0  3
```

We observe that the first record for each person has the same date of start (that is, value of `per`).

We then show the plot of the states and transitions but only for the follow-up starting at end of intervention (that is the date from which the groups are on identical regimens):

```
> par( mfrow=c(1,2) )
> for( whl in levels(P1$allocation) )
+ {
+ boxes( subset( P1, allocation==whl ),
+        boxpos=list(x=c(rep(20,4),rep(80,8)),
+                    y=c( seq(90,10,,4),
+                         rep(seq(92,8,,4),each=2)+rep(c(1,-1)*4,4) ) ),
+      # pos.arr=rep(c(5,4,6)/10,5)[-13][-15+i],
+        show.BE="nz", DR.sep=c(" (", ")"), eq.ht=FALSE, hmult=1.3,
+        scale.R=1000)
+ text(20,par("usr")[4],whl,adj=c(0.5,1),font=2,cex=2)
+ }
```

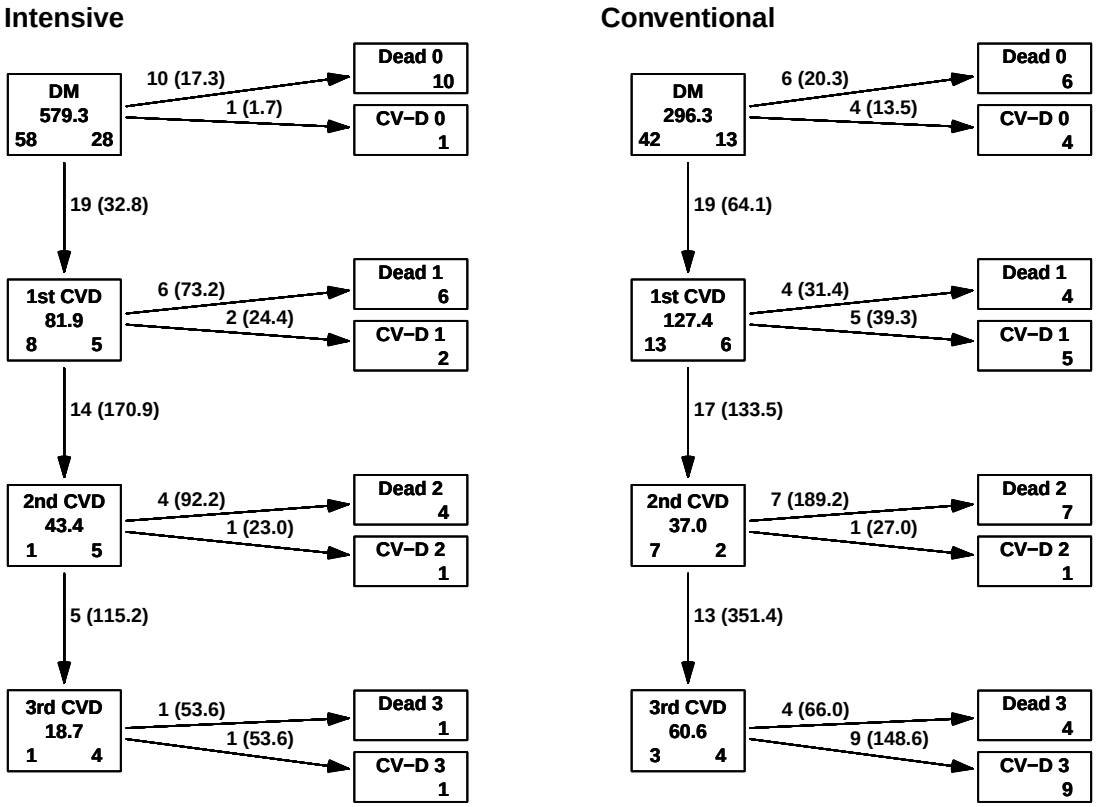

Figure 1.4: *Flow of patients from the Steno2 study by randomization group, restricted to the period **after** end of intervention (May 2001).*
The numbers in the middle of the boxes are the person-years, the number at the bottom of the boxes are the number of persons starting (left), resp. ending (right) their follow-up in each state. The numbers on the arrows are the number of transitions (transition-rates per 1000 PY).

Note that the dataset we have created here is only made for illustrational purposes, and will not be used in any analyses; there is no use for the date at which the two groups were merged as the effect of merging the group should be expected to be smooth and and is thus automatically modeled by non-linear effects of time since baseline.

1.4 Datasets saved for analysis

Finally we save the relevant datasets, namely `L0` (only one record of total follow-up for each person), and `L1` where follow-up is cut by CVD event. In both datasets we have a subdivision of

```
> save( st2, file="../data/st2.Rda" )
> save( L0, L1, file="../data/Lx.Rda")
```

1.5 Datsets saved for public use

In order to give the public access to the data from the trial we anonymize data by adding a randon quantity between −7 and +7 days to all dates recorded. This means that no dates or combinations of dates in the publicly available datasets occur in any real individual. Moreover we randomly assign new ID numbers to persons, so persons now have IDs in the range 1001–1160.

1.5.1 Merge to one data set

```
> clear()
> load( "../data/mic.Rda" )
> load( "../data/st2.Rda" )
> ( whd <- match(c("doEye","doPhC","doBlind"),names(mic)) )
[1] 7 8 9
> dmic <- aggregate( mic[,whd],
+                    mic["ID",drop=FALSE],
+                    FUN=function(x) mean(x,na.rm=TRUE) )
> str( dmic )
'data.frame':        160 obs. of  4 variables:
 $ ID     : num  1 2 3 4 5 6 7 8 9 10 ...
 $ doEye  : num  2015 2015 NaN NaN NaN ...
 $ doPhC  : num  2015 2008 NaN NaN NaN ...
 $ doBlind: num  NaN NaN NaN NaN NaN ...
> st2 <- merge( st2, dmic )
> # Locate all date variables
> names(st2)[dn <- grep( "do", names(st2) )]
 [1] "doBth"   "doDM"    "doBase"  "doCVD1"  "doCVD2"  "doCVD3"  "doESRD"  "doEnd"
 [9] "doDth"   "doEye"   "doPhC"   "doBlind"
> # Add +/- 7 days to the dates
> for( i in dn ) st2[,i] <- st2[,i] + runif(160,-7,7)/365
> # Check and fix the internal consistecy of dates:
> # doBth < doDM < doBase < doCVD1 < doCVD2 < doCVD3 < doDth
> # doBase < (doEye / doPhC / doBlind / doESRD) <doEnd < doDth
> # doBase < (=for visit1) doVis < (=for last visit) doEnd
>
> # The variables not constant within persons
> ( whm <- c( "ID", setdiff(names(mic), names(st2) ) ) )
[1] "ID"        "doVis"     "visit"     "aer_level" "med_aer"   "eds_l"     "eds_r"
[8] "autoprog"  "periprog"
> tot <- merge( mic[,whm], st2, all.x=TRUE )
```

Generate bogus ids:

```
> ID <- unique(tot$ID)
> ID <- ID[sample(1:length(ID))]
> id <- data.frame( ID=ID, id=1000+1:160 )
> dim( id )
[1] 160   2
> dim( tot )
[1] 748  25
> tot <- merge( id, tot, all=TRUE )[,-1]
> visits <- tot[,c("id","allocation","sex","visit","doVis",
+                  "aer_level","med_aer","eds_l","eds_r","autoprog","periprog",
+                  "doEye","doPhC","doBlind","baseCVD","deathCVD",
+                  "doBth","doDM","doBase","doCVD1","doCVD2","doCVD3",
+                  "doESRD","doEnd","doDth")]
> base <- aggregate( visits[,c("baseCVD","deathCVD",
```

```
+                              "doBth","doDM","doBase","doCVD1","doCVD2","doCVD3",
+                              "doESRD","doEnd","doDth")],
+                    visits[,c("id","allocation","sex")],
+                    FUN=function(x) mean(x,na.rm=TRUE) )
> names( visits )
 [1] "id"         "allocation" "sex"        "visit"      "doVis"      "aer_level"
 [7] "med_aer"    "eds_l"      "eds_r"      "autoprog"   "periprog"   "doEye"
[13] "doPhC"      "doBlind"    "baseCVD"    "deathCVD"   "doBth"      "doDM"
[19] "doBase"     "doCVD1"     "doCVD2"     "doCVD3"     "doESRD"     "doEnd"
[25] "doDth"
> names( base )
 [1] "id"         "allocation" "sex"        "baseCVD"    "deathCVD"   "doBth"
 [7] "doDM"       "doBase"     "doCVD1"     "doCVD2"     "doCVD3"     "doESRD"
[13] "doEnd"      "doDth"
> package.skeleton( "St2", c("visits","base"), path="../data/", force=TRUE )
```

Chapter 2

Statistical analysis of rates

2.1 Concepts

2.1.1 Outcomes

There are two broad categories of events:

- Deaths:
  - CVD-death (defined by *cause*)
  - Other
  - All (the union of the two previous)
- CVD-events:
  - First event
  - Any event
  - Number of events
- Life lost, quality years lived:

  These measures are derivatives of transition intensities (transition rates), basically summaries of the time spent in different states — the simplest being in the "alive" state, the more complicated being "without CVD", say. Since these are *absolute measures* and not relative measures, they are not easy to generalize from the patient population at hand, and so they are better modelled also including age, sex and other demographic characteristics, nd reported for select values of these.

2.1.2 Predictors

Beyond the randomization class (Intensive / Conventional), there is a specific interest in describing how mortality changes after intervention end, where the control group were "lifted" to the level of the Intensive group.

Hence we would primarily be interested in the effect of calendar time (or time since baseline which is effectively the same) on the mortality and in particular interaction between Intensive group and calendar time. This is best addressed by a parametric model for the mortality as function of time.

It is not likely that we will have power in the study to address effects of age and diabetes duration in any detail. Hence analysis of rates (mortality or CVD) will be by allocation group and time since randomization, and possibly bt CVD status (such as current number of CVD events).

2.1.3 Multistate models

The overview of the patient flow shown in the precious chapter lends itself to quantification of years lived in various states, by using estimated transition rates for all transitions to predict the probability of being in any of the states at any time, and subsequently quantifying how long patients in different groups have spent in these states. Typically, the intervention effect can be assessed on this scale as well by subtracting the years lived for patients in the two groups, but this difference is not likely to be generalizable beyond the specific patient group at hand, so we may consider modeling the rates using at least age and sex as predictors too, so that the intervention effect of years lived can be quantified as a function of these.

2.2 Age- and sex-distribution

The age-distribution at baseline is a bit ragged, and the sex-distribution is very skew (41 women, 119 men) as seen in figure 2.1:

```
> library( Epi )
> clear()
> load( file="../data/Lx.Rda" )
> lls()
  name mode class            size
1 L0   list Lexis data.frame 160 23
2 L1   list Lexis data.frame 318 23
> par( mfrow=c(1,2) )
> for( sx in levels(L0$sex)[2:1] )
+ {
+ with( subset(L0,sex==sx),
+       hist( age, col=if(sx=="M") "blue" else "red",
+               border=if(sx=="M") "blue" else "red",
+             breaks=seq(36,69,3), ylim=c(0,22),
+             main="", xlab="" ) )
+ par( new=TRUE )
+ with( subset(L0,sex==sx & allocation=="Intensive"),
+       hist( age, col="#AAAAAAAA",
+               border="transparent",
+             breaks=seq(36,69,3), ylim=c(0,22),
+             main="", xlab="" ) )
+ }
> mtext( "Age at baseline", side=1, line=-3, outer=TRUE )
```

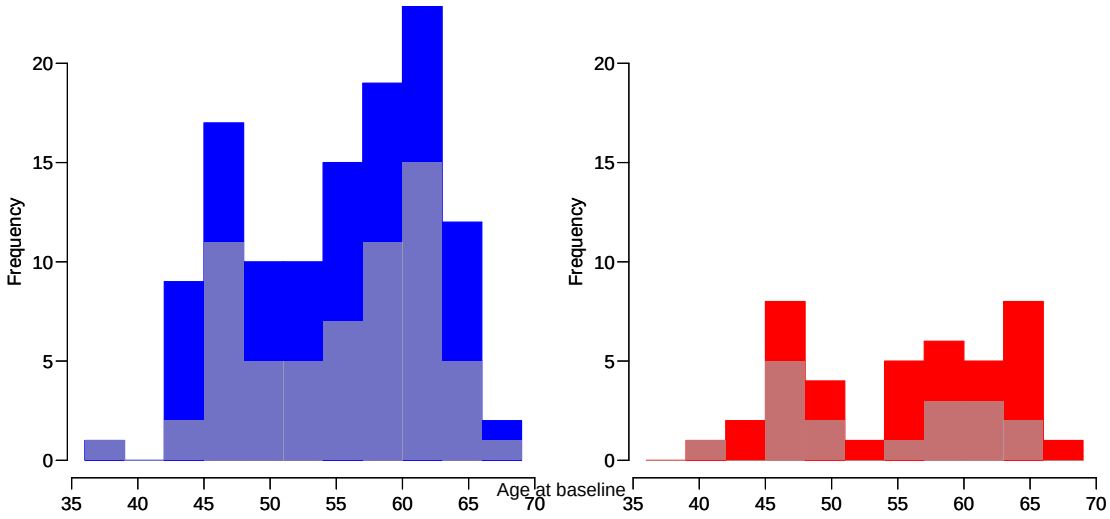

Figure 2.1: *Age-distribution among men (blue) and women (red), the light areas are the persons allocated to the intensive group. Age-classes are 3 years wide.*

Chapter 3

Mortality rates

3.1 Differences in median survival

First we compute the crude Kaplan-Meier-estimates of overall survival and CVD-free survival

To this end we take the previous analysis dataset but merge the two types of death:

```
> library( survival )
> library( Epi )
> clear()
> sessionInfo()
R version 3.3.0 (2016-05-03)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 14.04.4 LTS

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C               LC_TIME=en_US.UTF-8
 [4] LC_COLLATE=en_US.UTF-8     LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8       LC_NAME=C                  LC_ADDRESS=C
[10] LC_TELEPHONE=C             LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

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

other attached packages:
[1] Epi_2.5         survival_2.39-4

loaded via a namespace (and not attached):
 [1] cmprsk_2.2-7      MASS_7.3-44       plyr_1.8.3        Matrix_1.2-1
 [5] parallel_3.3.0    etm_0.6-2         Rcpp_0.11.6       splines_3.3.0
 [9] grid_3.3.0        numDeriv_2014.2-1 lattice_0.20-31
> load( file="../data/Lx.Rda" )
> lls()
  name mode class            size
1 L0   list Lexis data.frame 160 22
2 L1   list Lexis data.frame 318 22
> levels( L0 )
[1] "DM"   "Dead" "CV-D"
> levels( L1 )
 [1] "DM"      "1st CVD" "2nd CVD" "3rd CVD" "Dead 0"  "CV-D 0"  "Dead 1"  "CV-D 1"
 [9] "Dead 2"  "CV-D 2"  "Dead 3"  "CV-D 3"
> L1 <- Relevel( L1, list("Dead(no CVD)"=5:6,
+                         "Dead(1 CVD)"=7:8,
+                         "Dead(2 CVD)"=9:10,
+                         "Dead(3 CVD)"=11:12), first=FALSE )
```

```
      type     old          new
1  lex.Cst      DM           DM
2  lex.Cst 1st CVD      1st CVD
3  lex.Cst 2nd CVD      2nd CVD
4  lex.Cst 3rd CVD      3rd CVD
5  lex.Cst  Dead 0
6  lex.Cst  CV-D 0
7  lex.Cst  Dead 1
8  lex.Cst  CV-D 1
9  lex.Cst  Dead 2
10 lex.Cst  CV-D 2
11 lex.Cst  Dead 3
12 lex.Cst  CV-D 3
13 lex.Xst      DM           DM
14 lex.Xst 1st CVD      1st CVD
15 lex.Xst 2nd CVD      2nd CVD
16 lex.Xst 3rd CVD      3rd CVD
17 lex.Xst  Dead 0 Dead(no CVD)
18 lex.Xst  CV-D 0 Dead(no CVD)
19 lex.Xst  Dead 1  Dead(1 CVD)
20 lex.Xst  CV-D 1  Dead(1 CVD)
21 lex.Xst  Dead 2  Dead(2 CVD)
22 lex.Xst  CV-D 2  Dead(2 CVD)
23 lex.Xst  Dead 3  Dead(3 CVD)
24 lex.Xst  CV-D 3  Dead(3 CVD)
> levels( L1 )
[1] "DM"           "1st CVD"      "2nd CVD"      "3rd CVD"      "Dead(no CVD)"
[6] "Dead(1 CVD)"  "Dead(2 CVD)"  "Dead(3 CVD)"
> summary( L1 )
Transitions:
     To
From      DM 1st CVD 2nd CVD 3rd CVD Dead(no CVD) Dead(1 CVD) Dead(2 CVD) Dead(3 CVD)
  DM      41      86       0       0           33          0          0          0
  1st CVD  0      11      48       0            0         27          0          0
  2nd CVD  0       0       7      24            0          0         16          0
  3rd CVD  0       0       0       8            0          0          0         17
  Sum     41      97      55      32           33         27         16         17

Transitions:
     To
From       Records:  Events: Risk time:  Persons:
  DM            160      119    1870.64       160
  1st CVD        86       75     342.61        86
  2nd CVD        47       40     112.27        47
  3rd CVD        25       17      92.08        25
  Sum           318      251    2417.60       160
```

The Kaplan-Meier curve for overall death is based on all risk time in any fo the alive states, and the event is any death, whereas the event "death or CVD" is based only on risk time in "DM" and events being either "1st CVD" or "D0":

```
> ( Xall <- levels(L0)[2:3] )
[1] "Dead" "CV-D"
> ( Ccvd <- levels(L1)[1] )
[1] "DM"
> ( Xcvd <- levels(L1)[c(2,5)] )
[1] "1st CVD"      "Dead(no CVD)"
```

With this fixed we can set up the dataset needed for specification of the combined enpoint of death and first CVD (CVD-free survival), and then the two Kaplan-Meier analyses:

```
> Lcvd <- subset( L1, lex.Cst %in% Ccvd )
> # The ids are all in the same order
> all( L0$lex.id == Lcvd$lex.id )
[1] TRUE
> KMall <- survfit( Surv(lex.dur,lex.Xst %in% Xall) ~ allocation,
+                   data = L0 )
> KMcvd <- survfit( Surv(lex.dur,lex.Xst %in% Xcvd) ~ allocation,
+                   data = Lcvd )
> RTall <- do.call( cbind, summary( KMall, times=0:4*5)[c("time","n.risk")] )
> RTcvd <- do.call( cbind, summary( KMcvd, times=0:4*5 )[c("time","n.risk")] )
```

We can plot the two sets of survival curves:

```
> nt <- nrow(RTall)/2
> clr <- c("forestgreen","red")
> par( mfrow=c(1,2), mar=c(6,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> plot( KMall, conf.int=TRUE, col=clr, lwd=1, yaxs="i", ylim=c(0,1.01),
+       ylab="Overall survival" )
> abline( h=0.5, v=summary(KMall)$table[,"median"] )
> axis( side=1, at=1:21, labels=NA, tcl=-0.3 )
> axis( side=2, at=1:9/10, labels=NA, tcl=-0.3 )
> text( rep(1,2), c(9,5)/100, c("Intensive","Conventional"),
+       col=clr, font=2, cex=1.2, adj=0 )
> lines( KMall, col=clr, lwd=4 )
> for(j in 0:1 ) mtext( RTall[1:nt+nt*j,2], at= RTall[1:nt+nt*j,1],
+                       line=2.5+j, side=1, col=clr[j+1], font=2, cex=1.2 )
> plot( KMcvd, conf.int=TRUE, col=clr, lwd=1, yaxs="i", ylim=c(0,1.01),
+       ylab="CVD-free survival" )
> abline( h=0.5, v=summary(KMcvd)$table[,"median"] )
> axis( side=1, at=1:21, labels=NA, tcl=-0.3 )
> axis( side=2, at=1:9/10, labels=NA, tcl=-0.3 )
> lines( KMcvd, col=clr, lwd=4 )
> for(j in 0:1 ) mtext( RTcvd[1:nt+nt*j,2], at= RTcvd[1:nt+nt*j,1],
+                       line=2.5+j, side=1, col=clr[j+1], font=2, cex=1.2 )
```

We can make the same plot the other wau round

```
> nt <- nrow(RTall)/2
> clr <- c("forestgreen","red")
> par( mfrow=c(1,2), mar=c(6,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> plot( KMall, conf.int=TRUE, col=clr, lwd=1,
+       yaxs="i", yaxt="n", ylim=c(1,0),
+       ylab="Cumulative mortality" )
> abline( h=0.5, v=summary(KMall)$table[,"median"] )
> axis( side=1, at=1:21, labels=NA, tcl=-0.3 )
> axis( side=2, at=0:5/5, labels=5:0/5 )
> axis( side=2, at=1:9/10, labels=NA, tcl=-0.3 )
> text( rep(1,2), c(9,5)/100, c("Intensive","Conventional"),
+       col=clr, font=2, cex=1.2, adj=0 )
> lines( KMall, col=clr, lwd=4 )
> for(j in 0:1 ) mtext( RTall[1:nt+nt*j,2], at= RTall[1:nt+nt*j,1],
+                       line=3.5-j, side=1, col=clr[j+1], font=2, cex=1.2 )
> plot( KMcvd, conf.int=TRUE, col=clr, lwd=1,
+       yaxs="i", yaxt="n", ylim=c(1:0),
+       ylab="Death or CVD event" )
> abline( h=0.5, v=summary(KMcvd)$table[,"median"] )
> axis( side=1, at=1:21, labels=NA, tcl=-0.3 )
> axis( side=2, at=0:5/5, labels=5:0/5 )
> axis( side=2, at=1:9/10, labels=NA, tcl=-0.3 )
> lines( KMcvd, col=clr, lwd=4 )
> for(j in 0:1 ) mtext( RTcvd[1:nt+nt*j,2], at= RTcvd[1:nt+nt*j,1],
+                       line=3.5-j, side=1, col=clr[j+1], font=2, cex=1.2 )
```

The `summary.survfit` gives the median survival as well as the confidence intervals for these:

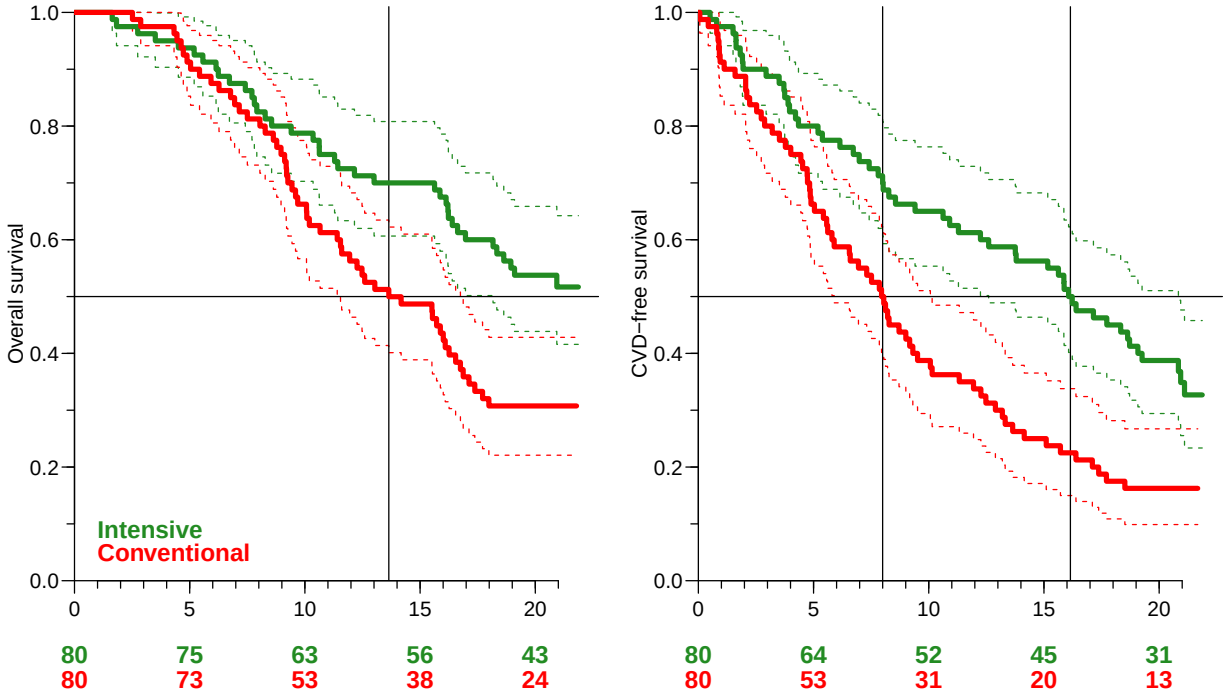

Figure 3.1: *Kaplan-Meier estimators for overall and CVD-free survival, numbers at the bottom is the number at risk at times 0, 5, etc.*

```
> round( summary( KMall )$table[,-(5:6)], 1 )
                        records n.max n.start events median 0.95LCL 0.95UCL
allocation=Intensive         80    80      80     38     NA    18.2      NA
allocation=Conventional      80    80      80     55   13.6    11.5    16.8
> round( summary( KMcvd )$table[,-(5:6)], 1 )
                        records n.max n.start events median 0.95LCL 0.95UCL
allocation=Intensive         80    80      80     52   16.1    12.6    20.8
allocation=Conventional      80    80      80     67    8.0     5.9    10.1
```

However we would like to see the confidence intervals for the *difference* between the groups. This is done by a simple bootstrap of the original dataset; we devise a function that computes the median survival in the two groups for the two types of outcome, for a resampled dataset and we then run this function 5000 times to get a bootstrap sample from the distribution of median survival times:

```
> Difmed <-
+ function()
+ {
+ wh <- sample(1:160,160,replace=TRUE)
+ KMall <- survfit( Surv(tsb,tsb+lex.dur,lex.Xst %in% Xall) ~ allocation,
+                   data = L0[wh,] )
+ KMcvd <- survfit( Surv(tsb,tsb+lex.dur,lex.Xst %in% Xcvd) ~ allocation,
+                   data = Lcvd[wh,] )
+ rbind( summary( KMall )$table["median"],
+        summary( KMcvd )$table["median"] )
+ }
```

We set up an array to hold the values of the median survival time for the two groups, and separately for the two types of outcome.

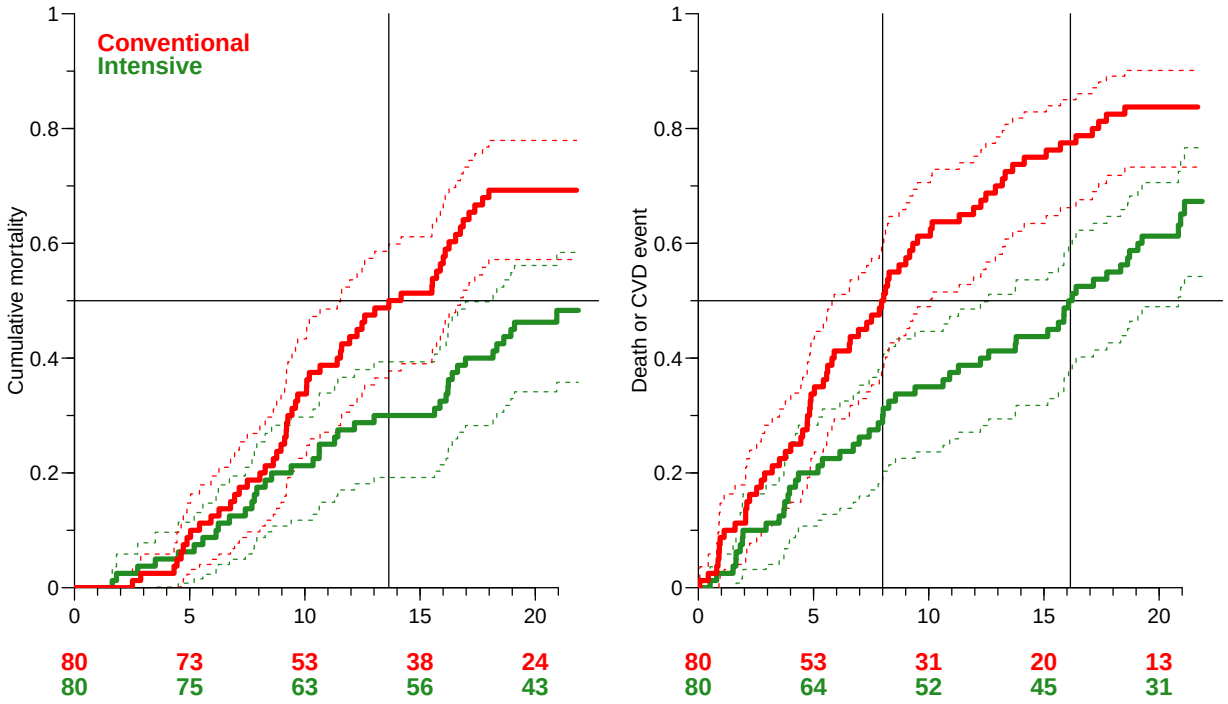

Figure 3.2: *Kaplan-Meier estimators for overall and CVD-free survival, numbers at the bottom is the number at risk at times 0, 5, etc.*

```
> meds <- NArray( list( 1:5000,
+                       type = c("All","CVD"),
+                 allocation = c("Int","Conv") ) )
> set.seed( 57976895 )
> system.time( for( i in 1:dim(meds)[1] ) meds[i,,] <- Difmed() )
   user  system elapsed
 73.997   0.004  73.998

> org <- meds
> meds <- org
> round( apply( meds, 2:3, function(x) mean(is.na(x) ) )*100, 2 )
     allocation
type    Int Conv
  All 60.80    0
  CVD  0.14    0
```

We see that there quite a substantial fraction of the simulaions that produce a missing median survival time for the intensive group for all-cause mortality. If we replace these with 21, we get an under-estimate of the median survival, and since the single imputation is mainly for one group also of the difference in median survival times:

```
> meds[is.na(meds)] <- 21
> meds <- meds[,,c(1,2,2)]
> dimnames(meds)[[3]][3] <- "I - C"
> dimnames(meds)[-1]
$type
[1] "All" "CVD"

$allocation
[1] "Int"   "Conv"  "I - C"
```

```
> meds[,,3] <- meds[,,"Int"]-meds[,,"Conv"]
> round( ftable( apply( meds,
+                       2:3,
+                       quantile,
+                       probs=c(0.5,0.025,0.975) ),
+                col.vars=c(3,1) ), 1 )
```

```
     allocation  Int            Conv            I - C
                 50% 2.5% 97.5%  50% 2.5% 97.5%   50% 2.5% 97.5%
type
All             21.0 16.8  21.0 13.6 10.8  16.5   6.4  2.2   9.6
CVD             16.1 12.2  19.7  8.0  5.8   9.5   8.3  4.0  12.6
```

Thus we see that the difference in median survival is (at least) 6.4 years (2.2;9.6), whereas the difference in median CVD-free survival is

3.2 All-cause mortality

First we model all-cause mortality as a function of time since randomization separately in the two groups with the intent of inspecting how mortality between the two groups develop.

In terms of the multistate figures in the previous chapter, this analysis corresponds to pooling of all alive states and pooling all dead states, and just analyzing transitions from alive to dead.

3.2.1 Splitting the dataset by time

First we split the dataset in intervals of one month (that is 365.25/12 days) in order to model mortality by time from baseline as a smooth function of time. This is because we shall assume that mortality varies continuously with time since baseline.

```
> library( survival )
> library( Epi )
> library( splines )
> clear()
> print( sessionInfo(), l=F )
```

```
R version 3.3.0 (2016-05-03)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 14.04.4 LTS

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

other attached packages:
[1] Epi_2.5         survival_2.39-4

loaded via a namespace (and not attached):
 [1] cmprsk_2.2-7      MASS_7.3-44       plyr_1.8.3        Matrix_1.2-1
 [5] parallel_3.3.0    etm_0.6-2         Rcpp_0.11.6       grid_3.3.0
 [9] numDeriv_2014.2-1 lattice_0.20-31
```

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

```
  name mode class            size
1 L0   list Lexis data.frame 160 22
2 L1   list Lexis data.frame 318 22
```

```
> S1 <- splitLexis( L1, breaks=seq(0,25,1/12), time.scale="tsb" )
> summary( S1 )
```

```
Transitions:
     To
From          DM 1st CVD 2nd CVD 3rd CVD Dead 0 CV-D 0 Dead 1 CV-D 1 Dead 2 CV-D 2 Dead 3
  DM       22407      86       0       0     24      9      0      0      0      0      0
  1st CVD      0    4120      48       0      0      0     14     13      0      0      0
  2nd CVD      0       0    1358      24      0      0      0      0     12      4      0
  3rd CVD      0       0       0    1114      0      0      0      0      0      0      5
  Sum      22407    4206    1406    1138     24      9     14     13     12      4      5

Transitions:
     To
From       CV-D 3  Records:  Events: Risk time:  Persons:
  DM            0     22526      119    1870.64       160
  1st CVD       0      4195       75     342.61        86
  2nd CVD       0      1398       40     112.27        47
  3rd CVD      12      1131       17      92.08        25
  Sum          12     29250      251    2417.60       160
```

Note that the events enumerated in the above summary are *not* the deaths; the outcome of interest in the analysis of all-cause mortality are the transitions to any of the death states:

```
> ( dst <- levels(L1)[-(1:4)] )
[1] "Dead 0" "CV-D 0" "Dead 1" "CV-D 1" "Dead 2" "CV-D 2" "Dead 3" "CV-D 3"
```

We start by modeling mortality by a fairly detailed function of time since baseline (`tsb`), and common linear effects of age and sex between the two groups. First we devise the knots for the `tsb`-effect:

```
> ( t.kn <- with( subset(S1,lex.Xst %in% dst),
+                 c(0,quantile( tsb+lex.dur, probs=(1:5-0.5)/5) )) )
      10%       30%       50%       70%       90%
 0.000000  4.648323  7.869678 10.193018 14.711020 17.091855
> m0 <- glm( (lex.Xst %in% dst) ~ allocation + Ns(tsb, kn=t.kn),
+            family = poisson,
+            offset = log(lex.dur),
+              data = S1 )
> mi  <- update( m0, . ~ . - Ns(tsb, kn=t.kn) +
+                  allocation:Ns(tsb, kn=t.kn) )
> ma  <- update( m0, . ~ . + I(age-tsb) + sex )
> mA  <- update( m0, . ~ . +   age      + sex )
> mia <- update( mi, . ~ . + I(age-tsb) + sex )
```

`m0` is the proportional hazards model, `mi` is the model where hazards varies separately for the two randomization groups, `ma` a proportional hazards model with effects of sex and (linear effect of) age at baseline (which, because the linear effect of time since baseline is in the model, will be the same as the effect of current age):

```
> rbind( ci.exp( ma, subset="age"),
+        ci.exp( mA, subset="age") )
             exp(Est.)     2.5%    97.5%
I(age - tsb)  1.109305 1.072611 1.147254
age           1.109305 1.072611 1.147254
> c( ma$deviance, mA$deviance )
[1] 1191.187 1191.187
```

With these models fitted we can now test for proportionality of hazards along the time since baseline with and without age and sex included, and we can test for the effect of age and sex with and without the proportional hazards assumption:

```
> zz <- as.matrix( anova( ma, m0, mi, mia, ma, test="Chisq" ) )
> rownames( zz ) <- c("","age+sex|ph","ph|NULL", "age+sex|!ph","ph|age+sex")
> round(zz,3)
            Resid. Df Resid. Dev Df Deviance Pr(>Chi)
                29241   1191.187 NA       NA       NA
age+sex|ph      29243   1234.257 -2  -43.070    0.000
ph|NULL         29238   1229.505  5    4.752    0.447
age+sex|!ph     29236   1186.012  2   43.492    0.000
ph|age+sex      29241   1191.187 -5   -5.174    0.395
```

We see that there is no evidence whatsoever of non-proportional hazards, so we can compute the estimated HR in the two different proportional hazards models, with and without control for age and sex:

```
> HR.IC <- rbind( ci.exp( m0, subset="allo", pval=TRUE ),
+                 ci.exp( ma, subset="allo", pval=TRUE ) )
> rownames(HR.IC)<- c("Allo|ph","Allo|ph+age+sex")
> round( cbind( HR.IC, 1/HR.IC[,c(1,3,2)] ), 3 )
               exp(Est.)  2.5% 97.5%     P exp(Est.) 97.5%  2.5%
Allo|ph            1.815 1.198 2.750 0.005     0.551 0.364 0.835
Allo|ph+age+sex    1.899 1.249 2.888 0.003     0.526 0.346 0.800
```

We put these estimates in the collect structure for all cuse mortality:

```
> load( file="../data/mainCI.Rda" )
> dimnames( mainCI )[[1]][1]
[1] "All cause mortality"

> zz <- HR.IC
> zz[,1:3] <- 1/HR.IC[,c(1,3,2)]
> mainCI[1,,] <- zz
> round( ftable( mainCI, row.vars=1 ), 3 )
                      model      raw                      age/sex
                      what  Estimate  2.5% 97.5%     P Estimate  2.5% 97.5%     P
outcome
All cause mortality             0.551 0.364 0.835 0.005    0.526 0.346 0.800 0.003
CVD mortality                   0.379 0.191 0.754 0.006    0.353 0.177 0.705 0.003
non-CVD mortality               0.703 0.413 1.197 0.195    0.686 0.401 1.172 0.168
Death or 1st CVD                   NA    NA    NA    NA       NA    NA    NA    NA
Death | CVD state               0.835 0.536 1.301 0.425    0.874 0.564 1.355 0.548
CVD event | CVD state           0.552 0.394 0.772 0.001    0.575 0.412 0.803 0.001
Retinopthy progression          0.668 0.507 0.881 0.004    0.673 0.511 0.887 0.005
Autonomic neuropathy            0.594 0.397 0.890 0.011    0.573 0.382 0.859 0.007
Peripheral neuropathy           1.120 0.707 1.774 0.630    1.101 0.694 1.747 0.683
Macroalbuminuria                0.516 0.316 0.842 0.008    0.495 0.302 0.811 0.005

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

Finally we can extract the estimates of the age and sex effects:

```
> as <- cbind(
+ ci.exp( ma , subset=c("age","sex"), pval=TRUE ),
+ ci.exp( mia, subset=c("age","sex"), pval=TRUE ) )
> colnames(as)[c(1,5)] <- c("effect|ph","  effect|!ph")
> round( as, 3 )
             effect|ph  2.5% 97.5%     P   effect|!ph  2.5% 97.5%     P
I(age - tsb)     1.109 1.073 1.147 0.000        1.110 1.073 1.148 0.000
sexM             1.477 0.898 2.428 0.124        1.493 0.908 2.456 0.114
```

We see that the effect of age and sex on mortality is pretty much the same as that in the general Danish population, with an annual increase in mortality by age in the vicinity of 10% and a male/female mortality ratio close to 1.5.

3.2.2 Danish population mortality

To substantiate this, we load the Danish population mortality data (available in the `Epi` package) and make a quick analysis of mortality in the age-range 45–100 in the period 1995–2012. It is well-known that population all-cause mortality increases exponentially by age after age 40 in most populations and that the M/F mortality RR is almost constant over the age-range, so we fit a simple model to the Danish population data to tease out the population-effects of age and sex:

```
> data( M.dk )
> summary( M.dk )
       A               sex             P               D                Y          
 Min.   : 0.00   Min.   :1.0   Min.   :1974   Min.   :   0.0   Min.   :   55.83  
 1st Qu.:24.75   1st Qu.:1.0   1st Qu.:1983   1st Qu.:  23.0   1st Qu.:18292.58  
 Median :49.50   Median :1.5   Median :1993   Median : 122.0   Median :29515.00  
 Mean   :49.50   Mean   :1.5   Mean   :1993   Mean   : 284.5   Mean   :26267.56  
 3rd Qu.:74.25   3rd Qu.:2.0   3rd Qu.:2003   3rd Qu.: 489.0   3rd Qu.:36592.42  
 Max.   :99.00   Max.   :2.0   Max.   :2012   Max.   :1276.0   Max.   :45979.33  
      rate        
 Min.   :  0.0000  
 1st Qu.:  0.6393  
 Median :  4.6487  
 Mean   : 52.3246  
 3rd Qu.: 44.2793  
 Max.   :717.9487  

> M.dk <- transform( subset( M.dk, A>44 & P>1994 ),
+                    sex=factor(sex,levels=2:1,labels=c("F","M")) )
> head( M.dk )

      A sex    P   D        Y     rate
3553 45   M 1995 127 37734.50 3.365620
3554 45   F 1995  85 37032.83 2.295261
3555 45   M 1996 128 37342.50 3.427730
3556 45   F 1996  82 36362.83 2.255050
3557 45   M 1997 114 37024.50 3.079042
3558 45   F 1997  92 35947.00 2.559323

> mp <- glm( D ~ I(A+0.5) + I(P-1994.5) + sex,
+            offset = log(Y/1000),
+            family = poisson,
+              data = M.dk )
> round( cbind( ci.exp( mp, Exp=F ),
+               ci.exp( mp, pval=T ) ), 4 )

              Estimate    2.5%   97.5% exp(Est.)   2.5%  97.5% P
(Intercept)    -3.7948 -3.8090 -3.7806    0.0225 0.0222 0.0228 0
I(A + 0.5)      0.0997  0.0995  0.0999    1.1048 1.1046 1.1050 0
I(P - 1994.5)  -0.0197 -0.0200 -0.0193    0.9805 0.9802 0.9809 0
sexM            0.3869  0.3829  0.3910    1.4725 1.4666 1.4784 0
```

showing that the annual increase in mortality by age is 10%, annual decrease in mortality by calendar time is 2% and the M/F mortality rate-ratio is 1.47. Specifically, the mortality (per 1000 PY) among men aged 55 in 1995 (that is born 1940) as a function of age is:

$$\mu_M(a) = \exp(-3.7948 + 0.0997 \times a - 0.0197 \times (a - 55) + 03869)$$

which we shall use to overlay in the display of the observed mortality rates in the Steno 2 population, by computing the population mortality for men in ages 55 and 75 in years 1995 and 2015 respectively — here computed two different ways:

```
> p.mort <- ci.exp( mp, ctr.mat=cbind( c(1, 1),
+                                      c(0,20)+55,
+                                      c(0,20),
+                                      c(1, 1)) )
> ci.pred( mp, data.frame(A =   55+c(0,20)-0.5, # The coding in the data is 0.5 years younger
+                         P = 1995+c(0,20)-0.5, # and 0.5 years earlier than the real time
+                       sex = "M",
+                         Y = 1000) )
   Estimate      2.5%     97.5%
1  7.962197  7.917715  8.006929
2 39.452057 39.251706 39.653430
> p.mort
     exp(Est.)      2.5%     97.5%
[1,]  7.962197  7.917715  8.006929
[2,] 39.452057 39.251706 39.653430
```

Since the modeling of the (log)population rates is linear in age and period we only need the rates at the endpoints 55 years, 1995 and 75 years, 2015 in order to be able to draw the predicted mortality rates as a function of time since baseline.

3.2.3 Mortality curves

We can now plot the two mortality curves and the HR between them; for the sex-age-controlled model we use men aged 55 to show the mortality rates (the majority of patients are men and the median age at baseline is 56.6 years):

```
> t.pt <- seq(0,20,0.1)
> CM <- Ns( t.pt, kn=t.kn )
> nd <- data.frame(tsb=t.pt,lex.dur=1000,age=55+t.pt,sex="M")
> rS.t <- ci.pred( mi, newdata=cbind(nd,allocation="Conventional") )
> rI.t <- ci.pred( mi, newdata=cbind(nd,allocation="Intensive") )
> rS.a <- ci.pred( mia,newdata=cbind(nd,allocation="Conventional") )
> rI.a <- ci.pred( mia,newdata=cbind(nd,allocation="Intensive") )
> pS.t <- ci.pred( m0, newdata=cbind(nd,allocation="Conventional") )
> pI.t <- ci.pred( m0, newdata=cbind(nd,allocation="Intensive") )
> pS.a <- ci.pred( ma, newdata=cbind(nd,allocation="Conventional") )
> pI.a <- ci.pred( ma, newdata=cbind(nd,allocation="Intensive") )
> # Check parameter sequencing using subset=
>  round( ci.exp( mi , subset=c("Intensive","Conventional") ), 3 )
                                          exp(Est.)  2.5%       97.5%
allocationIntensive:Ns(tsb, kn = t.kn)1       9.259 0.974      88.033
allocationIntensive:Ns(tsb, kn = t.kn)2       2.877 0.143      57.847
allocationIntensive:Ns(tsb, kn = t.kn)3       4.478 0.671      29.895
allocationIntensive:Ns(tsb, kn = t.kn)4      14.891 0.085    2603.902
allocationIntensive:Ns(tsb, kn = t.kn)5       3.044 0.967       9.582
allocationConventional                        0.136 0.002      10.014
allocationConventional:Ns(tsb, kn = t.kn)1    72.673 2.391    2209.022
allocationConventional:Ns(tsb, kn = t.kn)2    88.787 1.539    5121.615
allocationConventional:Ns(tsb, kn = t.kn)3    52.265 3.398     803.813
allocationConventional:Ns(tsb, kn = t.kn)4  2736.016 1.193 6274265.608
allocationConventional:Ns(tsb, kn = t.kn)5     5.225 1.569      17.397
> RR.t <- ci.exp( mi , subset=c("Intensive","Conventional"), ctr.mat=cbind(CM,-1,-CM) )
> RR.a <- ci.exp( mia, subset=c("Intensive","Conventional"), ctr.mat=cbind(CM,-1,-CM) )
```

Note that since we assume a proportional hazards model by age and sex, age at baseline and sex are not part of the RR calculations — the assumption is that the intervention effect is the same across all ages and for both sexes.

We now have the predicted mortality rates and the estimated HRs from the two models, so we can plot them side by side and see to what extent adjustment for age at baseline influences the estimates. First we define the colors and the corresponding transparent colors:

```
> clr <- c("forestgreen","red","black",gray(0.6))
> slr <- rgb( t( col2rgb(clr[1:2]) ), alpha=40, max=255 )
> par( mfrow=c(1,2), mar=c(0,0,0,0), oma=c(3,4,1,2), mgp=c(3,1,0)/1.6 )
> yl <- c(0.1,250)
> matplot( t.pt, cbind(rI.t,rS.t,RR.t),#pI.t,pS.t),
+          log="y", ylim=yl, xlab="",
+          type="l", lty=rep(c("21","63","63"),3), lend=1,
+          lwd=c(3,1,1), col=rep(clr[1:3],each=3) )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[1,2:3], col=clr[3], lty=1 )
> abline( h=1/HR.IC[1,  1], col=clr[3], lwd=3, lty=1 )
> matlines( t.pt, cbind(pI.t,pS.t),
+           type="l", lty=1, lwd=c(3,1,1), col=rep(clr[1:2],each=3) )
> matplot( t.pt, cbind(rI.a,rS.a,RR.a),#pI.a,pS.a),
+          log="y", ylim=yl, yaxt="n", ylab="",
+          type="l", lty=rep(c("21","63","63"),3), lend=1,
+          lwd=c(3,1,1), col=rep(clr[1:3],each=3) )
> matlines( t.pt, cbind(pI.a,pS.a),
+          type="l", lty=1, lwd=c(3,1,1), col=rep(clr[1:2],each=3) )
> lines( c(0,20), p.mort[,1], lwd=3, lty=1, col=clr[4] )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[2,2:3], col=clr[3], lty=1 )
> abline( h=1/HR.IC[2,  1], col=clr[3], lwd=3, lty=1 )
> mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
> text( rep(11,3), 10*0.70^c(2,1,3,4), c("Intensive","Conventional","Rate ratio","Population"),
+       col=clr, font=2, cex=1.2, adj=0 )
> mtext( "Mortality rate per 1000 PY / Hazard ratio Intensive vs Conventional",
+        side=2, line=3, las=0, outer=TRUE )
```

I order to plot the rightmost panel for the ESM together with the same for the CVD mortality we save the relevant structures:

```
> save( t.pt, HR.IC,
+             rI.t, rS.t,
+             rI.a, rS.a,
+             pI.t, pS.t, RR.t,
+             pI.a, pS.a, RR.a,
+       file = "../data/all-mort.Rda" )
```

From figure 3.3 it is seen that the intervention effect on overall mortality is not formally manifest till some 10 years after baseline, and there also seems to be a diminishing effect some 10 years after end of intervention. Comparison with the population rates shows that the intensive group has a mortality that is roughly mid-way between that of the conventional group and the population mortality.

Also, although there is no formal interaction between time since randomization and treatment group, when using the interaction model, the mortality in the intensive group tends to be larger in the first 2 years, but with very large confidence intervals for the HR.

3.2.4 DM duration

It is of some interest to assess to what extent diabetes duration has an effect om mortality, and in particular if the intervention as claimed by some is potentially harmful for persons with long duration of diabetes at entry. This amounts to investigation of the effect of duration and its interaction with allocation. Thus we expand the model for total mortality with an effect of duration, and further with duration terms separately for the two intervention groups, and finally with quadratic terms of duration for the two groups, to allow more flexible differences between duration effects in the two randomization groups.

```
> round( ci.exp(ma), 3 )
```

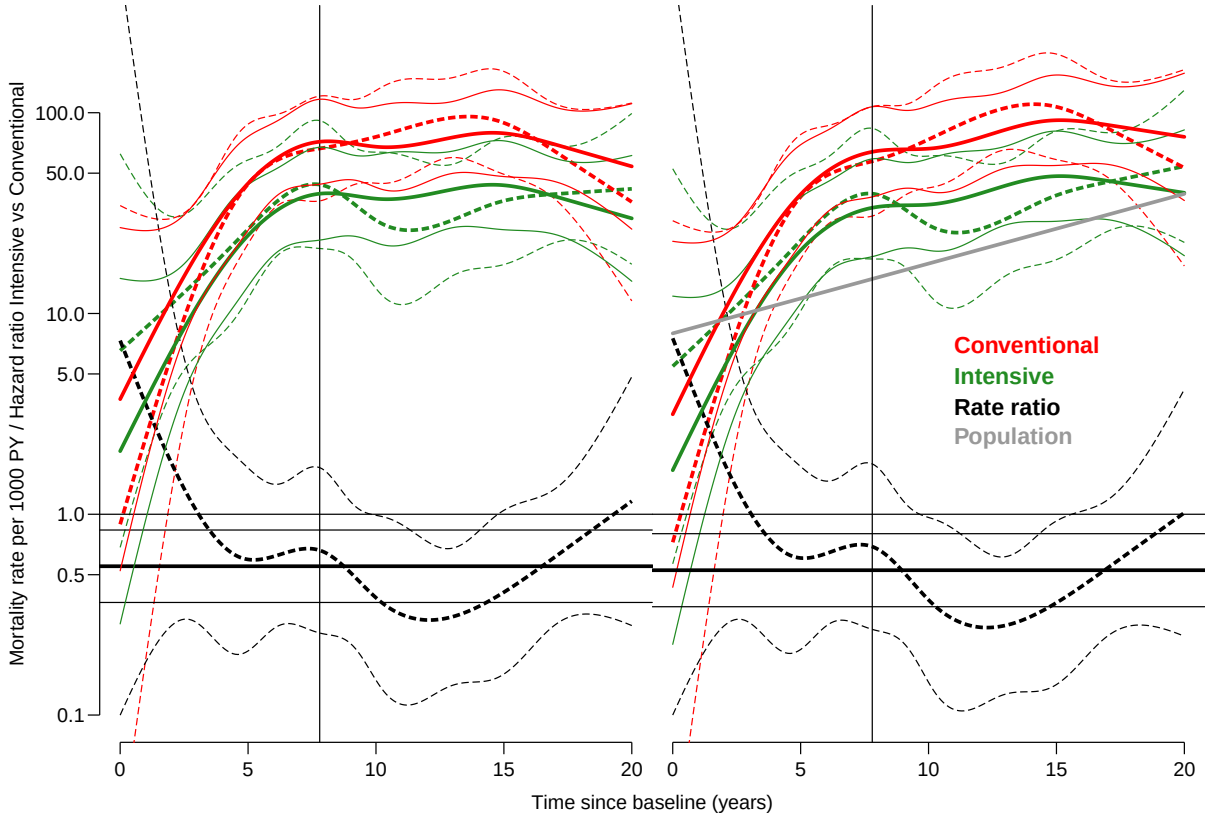

Figure 3.3: *Overall mortality rates and hazard ratio in the two groups. The full lines are hazards and HRs assuming proportional hazards, and the vertical line indicates the intervention end, broken lines are the hazards without the proportional hazards assumption. Left panel is models with only group and time since baseline, the models in the right panel also includes sex and age at baseline as a linear term on the log-mortality scale, and mortality rates are shown for a man aged 55 at baseline. The full black line is the population mortality for men aged 55 at start, aging to 75, over the period 1995–2015.*

```
                        exp(Est.)  2.5%     97.5%
(Intercept)                 0.000 0.000     0.000
allocationConventional      1.899 1.249     2.888
Ns(tsb, kn = t.kn)1        21.956 3.362   143.382
Ns(tsb, kn = t.kn)2        19.638 1.908   202.077
Ns(tsb, kn = t.kn)3        18.422 3.960    85.697
Ns(tsb, kn = t.kn)4       235.790 3.143 17691.556
Ns(tsb, kn = t.kn)5         5.720 2.573    12.717
I(age - tsb)                1.109 1.073     1.147
sexM                        1.477 0.898     2.428
> md  <- update( ma , . ~ . +           I( dur-tsb   ) )
> md1 <- update( ma , . ~ . + allocation:I( dur-tsb   ) )
> md2 <- update( md1, . ~ . + allocation:I((dur-tsb)^2) )
> anova( ma, md, md1, md2, test="Chisq" )[-1,-(1:2)]
  Df Deviance Pr(>Chi)
2  1  3.03918  0.08128
3  1  1.13584  0.28653
4  2  0.52607  0.76871
```

Thus we see that diabetes duration at baseline has a marginally significant effect (p=0.081), but no evidence of differential effect between the intensive and conventional groups. The estimated trend by duration, jointly and separately are here:

```
> round( ci.exp( md , subset=c("all","dur") ), 4 )
                        exp(Est.)   2.5%  97.5%
allocationConventional     1.7976 1.1772 2.7451
I(dur - tsb)               1.0342 0.9968 1.0731

> round( ci.exp( md1, subset=  "all"        ), 4 )

                                     exp(Est.)   2.5%  97.5%
allocationConventional                  2.4691 1.1949 5.1021
allocationIntensive:I(dur - tsb)        1.0594 1.0012 1.1209
allocationConventional:I(dur - tsb)     1.0189 0.9717 1.0683

> round( ( rbind( ci.exp( md , subset="dur" ),
+                 ci.exp( md1, subset="dur" ) )-1)*100, 1 )

                                     exp(Est.) 2.5% 97.5%
I(dur - tsb)                               3.4 -0.3   7.3
allocationIntensive:I(dur - tsb)           5.9  0.1  12.1
allocationConventional:I(dur - tsb)        1.9 -2.8   6.8
```

We see a borderline significant effect of duration (HR=3.4%/year (−0.3,7.3), p=0.081), but no significant difference in this between the two groups, p=0.286, although there is a very weak indication that DM duration at baseline has different effects in the two groups. In the intensive group there is an increase of mortality of 5.2%/year of DM duration, in the conventional group only 1.9%/year, not significantly different though.

It is however more informative to see the HR between intensive and conventional by duration as predicted by the two interaction models (the linear and the quadratic):

```
> CD <- cbind( -1, 0:20, -(0:20), (0:20)^2, -(0:20)^2 )
> rownames(CD) <- 0:20
> head( CD )
  [,1] [,2] [,3] [,4] [,5]
0   -1    0    0    0    0
1   -1    1   -1    1   -1
2   -1    2   -2    4   -4
3   -1    3   -3    9   -9
4   -1    4   -4   16  -16
5   -1    5   -5   25  -25

> round( cbind( RR1 <- ci.exp( md1, subset="all", ctr.mat=CD[,1:3], pval=TRUE ),
+               RRq <- ci.exp( md2, subset="all", ctr.mat=CD      , pval=TRUE ) ), 4 )

   exp(Est.)   2.5%  97.5%      P exp(Est.)   2.5%  97.5%      P
0     0.4050 0.1960 0.8369 0.0147    0.4428 0.1620 1.2102 0.1123
1     0.4211 0.2158 0.8218 0.0112    0.4451 0.1921 1.0312 0.0590
2     0.4379 0.2369 0.8094 0.0084    0.4491 0.2225 0.9066 0.0255
3     0.4553 0.2590 0.8003 0.0063    0.4548 0.2506 0.8255 0.0096
4     0.4734 0.2818 0.7953 0.0047    0.4623 0.2736 0.7812 0.0039
5     0.4922 0.3046 0.7954 0.0038    0.4718 0.2895 0.7689 0.0026
6     0.5118 0.3266 0.8020 0.0035    0.4832 0.2981 0.7832 0.0032
7     0.5321 0.3467 0.8166 0.0039    0.4968 0.3018 0.8178 0.0059
8     0.5533 0.3640 0.8410 0.0056    0.5126 0.3033 0.8665 0.0126
9     0.5753 0.3775 0.8766 0.0101    0.5310 0.3050 0.9246 0.0253
10    0.5981 0.3869 0.9247 0.0208    0.5521 0.3082 0.9888 0.0457
11    0.6219 0.3923 0.9860 0.0434    0.5762 0.3139 1.0575 0.0752
12    0.6467 0.3941 1.0610 0.0844    0.6036 0.3223 1.1304 0.1148
13    0.6724 0.3930 1.1502 0.1473    0.6346 0.3332 1.2089 0.1667
14    0.6991 0.3897 1.2542 0.2300    0.6698 0.3461 1.2962 0.2342
15    0.7269 0.3846 1.3737 0.3260    0.7096 0.3601 1.3983 0.3216
16    0.7558 0.3783 1.5098 0.4278    0.7546 0.3736 1.5242 0.4325
17    0.7859 0.3711 1.6640 0.5290    0.8055 0.3846 1.6870 0.5663
18    0.8171 0.3633 1.8378 0.6253    0.8630 0.3909 1.9053 0.7154
19    0.8496 0.3550 2.0333 0.7143    0.9281 0.3908 2.2043 0.8658
20    0.8834 0.3464 2.2525 0.7951    1.0019 0.3832 2.6193 0.9969
```

```
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> matplot( 0:20, cbind( RR1[,-4], RRq[,-4] ),
+          log="y", type="l", lwd=c(4,1,1), lty=1,
+          col=rep(c("black","gray"),each=3), ylim=c(0.2,2),
+          xlab="Diabetes duration at baseline",
+          ylab="Mortality HR: Intensive vs Conventional" )
> abline( h=1 )

> CD <- cbind( -1, 0:20, -(0:20), (0:20)^2, -(0:20)^2 )
> rownames(CD) <- 0:20
> head( CD )
  [,1] [,2] [,3] [,4] [,5]
0   -1    0    0    0    0
1   -1    1   -1    1   -1
2   -1    2   -2    4   -4
3   -1    3   -3    9   -9
4   -1    4   -4   16  -16
5   -1    5   -5   25  -25
> round( cbind( RR1 <- ci.exp( md1, subset="all", ctr.mat=CD[,1:3], pval=TRUE ),
+               RRq <- ci.exp( md2, subset="all", ctr.mat=CD      , pval=TRUE ) ), 4 )
   exp(Est.)   2.5%  97.5%      P exp(Est.)   2.5%  97.5%      P
0     0.4050 0.1960 0.8369 0.0147    0.4428 0.1620 1.2102 0.1123
1     0.4211 0.2158 0.8218 0.0112    0.4451 0.1921 1.0312 0.0590
2     0.4379 0.2369 0.8094 0.0084    0.4491 0.2225 0.9066 0.0255
3     0.4553 0.2590 0.8003 0.0063    0.4548 0.2506 0.8255 0.0096
4     0.4734 0.2818 0.7953 0.0047    0.4623 0.2736 0.7812 0.0039
5     0.4922 0.3046 0.7954 0.0038    0.4718 0.2895 0.7689 0.0026
6     0.5118 0.3266 0.8020 0.0035    0.4832 0.2981 0.7832 0.0032
7     0.5321 0.3467 0.8166 0.0039    0.4968 0.3018 0.8178 0.0059
8     0.5533 0.3640 0.8410 0.0056    0.5126 0.3033 0.8665 0.0126
9     0.5753 0.3775 0.8766 0.0101    0.5310 0.3050 0.9246 0.0253
10    0.5981 0.3869 0.9247 0.0208    0.5521 0.3082 0.9888 0.0457
11    0.6219 0.3923 0.9860 0.0434    0.5762 0.3139 1.0575 0.0752
12    0.6467 0.3941 1.0610 0.0844    0.6036 0.3223 1.1304 0.1148
13    0.6724 0.3930 1.1502 0.1473    0.6346 0.3332 1.2089 0.1667
14    0.6991 0.3897 1.2542 0.2300    0.6698 0.3461 1.2962 0.2342
15    0.7269 0.3846 1.3737 0.3260    0.7096 0.3601 1.3983 0.3216
16    0.7558 0.3783 1.5098 0.4278    0.7546 0.3736 1.5242 0.4325
17    0.7859 0.3711 1.6640 0.5290    0.8055 0.3846 1.6870 0.5663
18    0.8171 0.3633 1.8378 0.6253    0.8630 0.3909 1.9053 0.7154
19    0.8496 0.3550 2.0333 0.7143    0.9281 0.3908 2.2043 0.8658
20    0.8834 0.3464 2.2525 0.7951    1.0019 0.3832 2.6193 0.9969
> par( mfrow=c(1,2), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> hist( L0$dur, breaks=seq(0,32,2), col="black", main="",
+       xlab="Diabetes duration at baseline" )
> matplot( 0:20, cbind( RR1[,-4], RRq[,-4] ),
+          log="y", type="l", lwd=c(4,1,1), lty=1, col="black", ylim=c(0.2,2),
+          xlab="Diabetes duration at baseline",
+          ylab="HR: Intensive vs Conventional" )
> abline( h=1 )

> round( RR1[seq(1,21,5),], 2 )
   exp(Est.) 2.5% 97.5%    P
0       0.41 0.20  0.84 0.01
5       0.49 0.30  0.80 0.00
10      0.60 0.39  0.92 0.02
15      0.73 0.38  1.37 0.33
20      0.88 0.35  2.25 0.80
```

Thus the linear effect model corresponds to a HR between intensive and standard of 0.41 at baseline duration 0; 0.49 at 5 years; 0.60 at 10 years and 0.88 at 20 years - only significantly less than 1 for durations less than 10 years. Very similar effect is seen for the quadratic extension; figure **??**. In conclusion, there is no significant effect of duration at baseline and neither any sign of differential intervention effect by diabetes duration.

Hence we shall not explore this further.

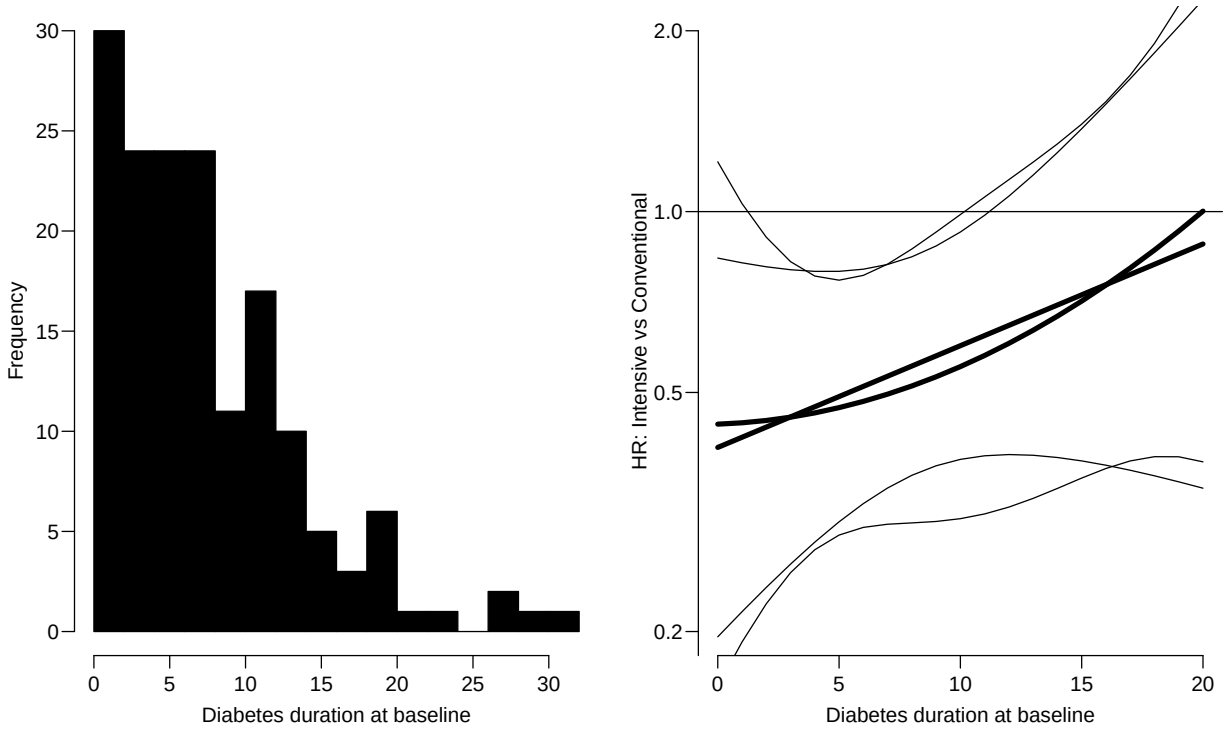

Figure 3.4: *Left: Distribution of duration of diabetes at baseline. Right: Hazard ratio between intensive and conventional by diabetes duration at baseline, estimated from two different interaction models, one with log-linear effects and another with log-quadratic effects of DM duration on HR.*

3.2.5 Survival

The analyses of mortality rates immediately translate into survival curves, the model with age-correction of course to survival curves for different ages at baseline — for the latter we shall use age 45,50,. . . ,65; corresponding approximately to the 10,30,50,70 and 90 age-percentiles at baseline, and of course give estimates separately for men and women:

```
> round( quantile( L0$age, probs=1:9/10 ), 1 )
 10%  20%  30%  40%  50%  60%  70%  80%  90%
45.4 47.4 50.0 54.1 56.6 57.9 60.4 62.2 64.2
> round( do.call( rbind,
+                 with( L0, tapply( age, sex,
+                                   quantile, probs=seq(1,9,2)/10 ) ) ), 1 )
   10%  30%  50%  70%  90%
F 46.1 50.3 56.8 60.7 64.6
M 45.3 50.0 56.6 60.1 63.5
```

We initially use the model `mi` from above; it is a model with a separate mortality for each group, modelled by a smooth function of time since baseline. Thus it is *not* a proportional hazards model; it is a model that basically models the mortality rates for each group separately, pretty much the machinery used to generate Kaplan-Meier curves for each group.

First we check that we get what we want from the `subset=` argument to `ci.exp`, and then we use it to compute the cumulative hazards with confidence bands

```
> ci.exp( mi, subset=c("cept","ensive") )
                                          exp(Est.)         2.5%        97.5%
(Intercept)                            0.006534735 0.0006839742    0.0624333
allocationIntensive:Ns(tsb, kn = t.kn)1  9.258952868 0.9738221988   88.0327110
allocationIntensive:Ns(tsb, kn = t.kn)2  2.876667150 0.1430540711   57.8467556
allocationIntensive:Ns(tsb, kn = t.kn)3  4.477914618 0.6707393510   29.8949499
allocationIntensive:Ns(tsb, kn = t.kn)4 14.891396861 0.0851620683 2603.9022416
allocationIntensive:Ns(tsb, kn = t.kn)5  3.043918250 0.9669421010    9.5822059
> ci.exp( mi, subset=c("cept","tional") )
                                             exp(Est.)         2.5%        97.5%
(Intercept)                             6.534735e-03 0.0006839742 6.243330e-02
allocationConventional                  1.364352e-01 0.0018588457 1.001404e+01
allocationConventional:Ns(tsb, kn = t.kn)1 7.267304e+01 2.3908186795 2.209022e+03
allocationConventional:Ns(tsb, kn = t.kn)2 8.878695e+01 1.5391867900 5.121615e+03
allocationConventional:Ns(tsb, kn = t.kn)3 5.226498e+01 3.3983404357 8.038125e+02
allocationConventional:Ns(tsb, kn = t.kn)4 2.736016e+03 1.1930935186 6.274266e+06
allocationConventional:Ns(tsb, kn = t.kn)5 5.225328e+00 1.5694436068 1.739728e+01
> ( intl <- mean(diff(t.pt)) )
[1] 0.1
> chI <- ci.cum( mi, subset=c("cept","ensive"), ctr.mat=cbind(1,  CM), intl=intl, ci.Exp=TRUE )
> chC <- ci.cum( mi, subset=c("cept","tional"), ctr.mat=cbind(1,1,CM), intl=intl, ci.Exp=TRUE )
> round( cbind(t.pt,chI,chC)[20+1:10,], 5 )
      t.pt Estimate    2.5%   97.5%     Erf Estimate    2.5%   97.5%     Erf
 [1,]  2.0  0.01810 0.00417 0.07858 4.34025  0.00590 0.00065 0.05371 9.10043
 [2,]  2.1  0.01925 0.00460 0.08056 4.18584  0.00657 0.00078 0.05541 8.42769
 [3,]  2.2  0.02042 0.00505 0.08248 4.03948  0.00731 0.00094 0.05708 7.81007
 [4,]  2.3  0.02162 0.00554 0.08435 3.90089  0.00811 0.00112 0.05875 7.24392
 [5,]  2.4  0.02286 0.00606 0.08619 3.76979  0.00898 0.00134 0.06041 6.72569
 [6,]  2.5  0.02414 0.00662 0.08800 3.64592  0.00993 0.00159 0.06209 6.25198
 [7,]  2.6  0.02545 0.00721 0.08980 3.52901  0.01096 0.00188 0.06380 5.81952
 [8,]  2.7  0.02679 0.00784 0.09159 3.41881  0.01208 0.00223 0.06554 5.42521
 [9,]  2.8  0.02817 0.00850 0.09340 3.31506  0.01329 0.00262 0.06735 5.06610
[10,]  2.9  0.02960 0.00920 0.09523 3.21751  0.01460 0.00308 0.06922 4.73945
```

We now have the cumulative hazards evaluated at the points `t.pt` for the two groups (with confidence intervals), and we can thus change them into survival curves with c.i. which we plot with shaded confidence bands:

```
> pls <-
+ function(yaxt="s",ci=TRUE,xl="Time since baseline (years)")
+ {
+ matplot( t.pt, exp( -cbind(chI[,1],chC[,1]) ),
+          col=clr[1:2], lty=1, lwd=4, type="l",
+          yaxt=yaxt, yaxs="i", ylim=c(0,1), ylab="",
+          xlim=c(0,20), xaxs="i", xlab=xl )
+ if(ci)
+   {
+   polygon( c(t.pt,rev(t.pt)), exp(-c(chI[,2],rev(chI[,3])) ), col=slr[1], border="transparent" )
+   polygon( c(t.pt,rev(t.pt)), exp(-c(chC[,2],rev(chC[,3])) ), col=slr[2], border="transparent" )
+   }
+ }
> par( mar=c(3,3,1,2), bty="n", mgp=c(3,1,0)/1.6, las=1 )
> pls()
> axis( side=4, at=0:5*20 /100)
> axis( side=4, at=0:10*10/100, tcl=-0.4, labels=NA )
> axis( side=4, at=0:20*5 /100, tcl=-0.3, labels=NA )
> axis( side=4, at=0:100  /100, tcl=-0.2, labels=NA )
> mtext( "Survival", side=2, las=0, line=2 )
```

We can compare the smoothed version of the survival curve with the Kaplan-Meier-estimates for the two groups, and we see that they are pretty much alike:

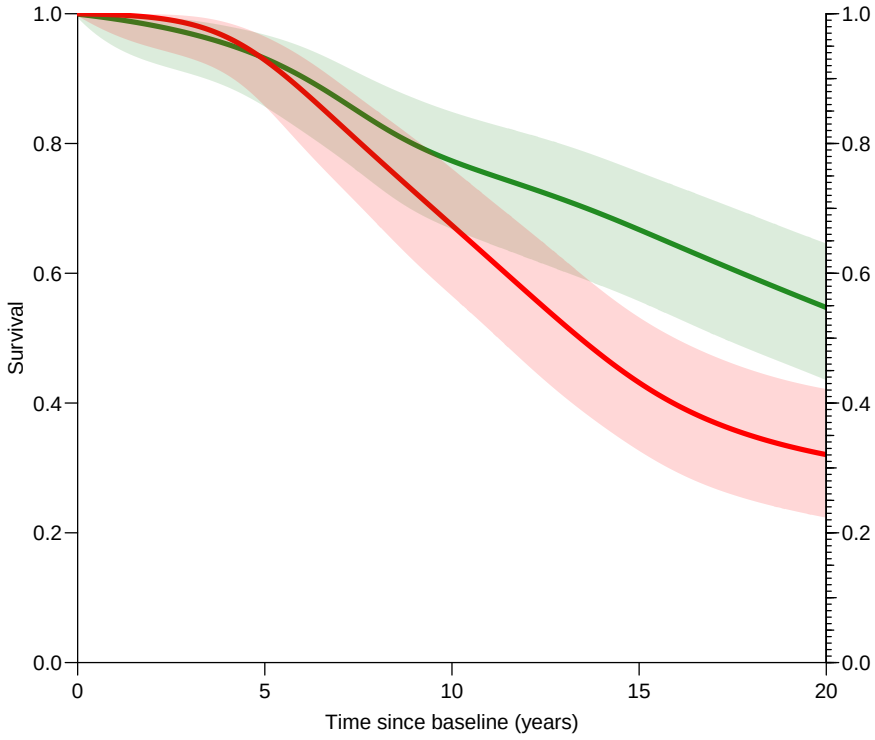

Figure 3.5: *Overall survival in the Steno 2 study. The smooth curves are from a model assuming a smoothly varying mortality rate by time since baseline.*

```
> par( mfrow=c(1,2), mar=c(3,0,1,1), oma=c(0,3,0,0), bty="n",
+     mgp=c(3,1,0)/1.6, las=1 )
> km <- survfit( Surv(tsb,tsb+lex.dur,lex.Xst %in% dst) ~ allocation, data=L1 )
> pls(ci=FALSE)
> mtext( "Survival", side=2, las=0, line=2 )
> abline( v=seq(0,20,5), h=0:10/10, col=gray(0.8) )
> par( new=TRUE )
> pls()
> pls(yaxt="n")
> lines(km,col=clr[1:2],mark.time=FALSE,conf.int=FALSE,lwd=2)
> lines(km,col=clr[1:2],mark.time=FALSE,conf.int="only",lwd=1,lty=3)
```

The curves in figure 3.6 refer to survival in a population of type 2 DM patients with an age-composition as the patient population in the Steno 2 study. The right hand panel has overlaid the traditional Kaplan-Meier curves that estimate the same quantities (the proportion of the patient population surviving to a given time) but clearly in a clinically much less credible form.

The difference between the intensive and standard groups is best summarized by the HR, either as constant or as time-varying as shown in figure 3.3. The *absolute* difference between regimens only makes sense if referred to a specific age and sex, because these two variables are such powerful predictors of mortality.

We therefore show the *estimated* survival for men and women entering the study at ages 45, 50, 55, 60 and 65 in order to make the absolute magnitudes of the survival probabilities interpretable. This will also make it possible to show the intervention effect relative to the general aging effect:

```
> # Array to collect estimated survival probabilities
> tabSurv <- NArray( list( grp = c("Int","Cnv"),
```

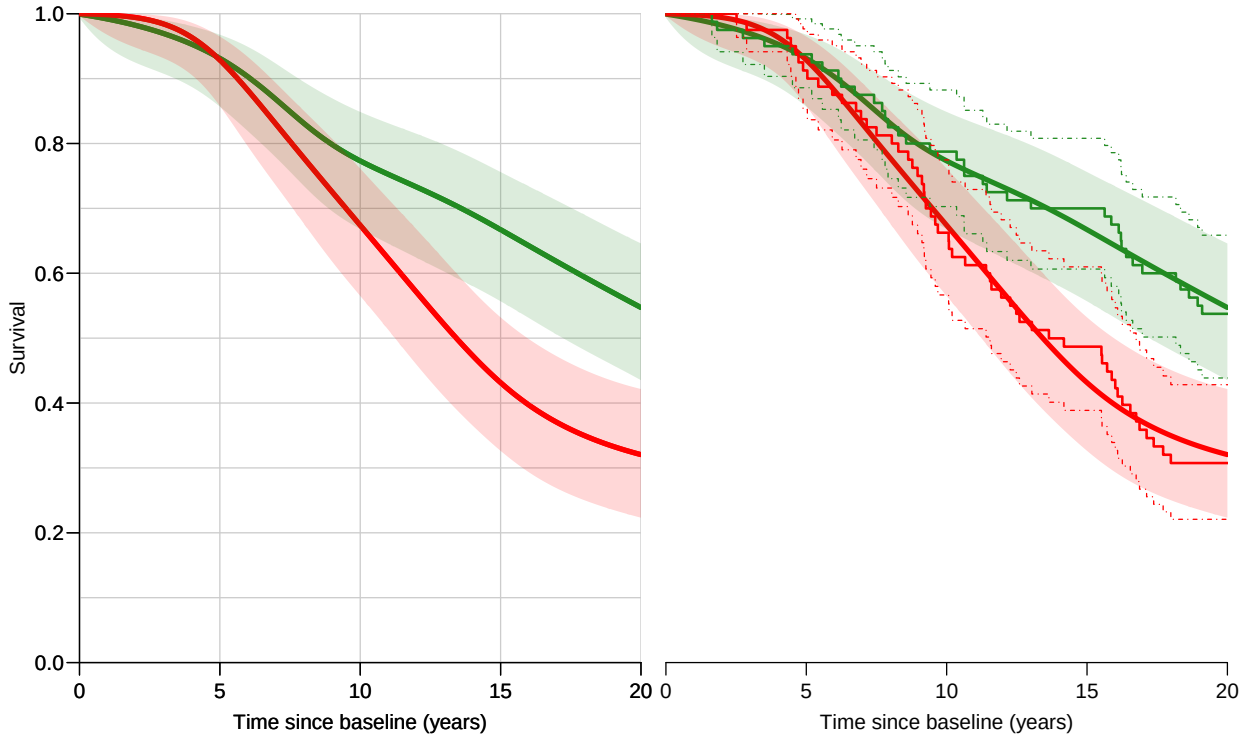

Figure 3.6: *Overall survival in the Steno 2 study. The smooth curves with shaded confidence bands (identical in the two panels) are from a model assuming a smoothly varying mortality rate by time since baseline, derived from the mortalities shown in the left panel of figure 3.3. The ragged curves overlaid in the right panel are the traditional Kaplan-Meier estimates, conveying a lot of unsubstantiated noise to the survival curves.*

```
+                          sex = c("M","F"),
+                          age = seq(45,65,5),
+                          tfr = t.pt,
+                         what = c("est","lo","up") ) )
> ci.exp( mia, subset=c("sex","age","cept","sive") )
                                              exp(Est.)         2.5%        97.5%
sexM                                       1.493121e+00 9.077945e-01 2.455854e+00
I(age - tsb)                               1.110024e+00 1.073230e+00 1.148079e+00
(Intercept)                                1.174012e-05 5.527660e-07 2.493467e-04
allocationIntensive:Ns(tsb, kn = t.kn)1 9.834423e+00 1.029889e+00 9.390903e+01
allocationIntensive:Ns(tsb, kn = t.kn)2 3.356022e+00 1.663345e-01 6.771227e+01
allocationIntensive:Ns(tsb, kn = t.kn)3 5.399008e+00 8.080748e-01 3.607251e+01
allocationIntensive:Ns(tsb, kn = t.kn)4 2.075319e+01 1.179995e-01 3.649973e+03
allocationIntensive:Ns(tsb, kn = t.kn)5 4.071501e+00 1.291699e+00 1.283357e+01
> ci.exp( mia, subset=c("sex","age","cept","onal") )
                                              exp(Est.)         2.5%        97.5%
sexM                                       1.493121e+00 9.077945e-01 2.455854e+00
I(age - tsb)                               1.110024e+00 1.073230e+00 1.148079e+00
(Intercept)                                1.174012e-05 5.527660e-07 2.493467e-04
allocationConventional                     1.328888e-01 1.765649e-03 1.000167e+01
allocationConventional:Ns(tsb, kn = t.kn)1 7.324541e+01 2.348622e+00 2.284271e+03
allocationConventional:Ns(tsb, kn = t.kn)2 1.143154e+02 1.931006e+00 6.767462e+03
allocationConventional:Ns(tsb, kn = t.kn)3 7.235271e+01 4.618208e+00 1.133538e+03
allocationConventional:Ns(tsb, kn = t.kn)4 4.615768e+03 1.904531e+00 1.118665e+07
allocationConventional:Ns(tsb, kn = t.kn)5 7.922665e+00 2.358872e+00 2.660959e+01
> par( mfrow=c(2,5), mar=c(3,0,1,1), oma=c(2,4,2,0) )
> for( sx in 1:0 )
+ for( a in seq(45,65,5) )
```

```
+    {
+ chI <- ci.cum( mia, subset=c("sex","age","cept","sive"),
+                 ctr.mat=cbind(sx,a,1,  CM), intl=intl, ci.Exp=TRUE )
+ chC <- ci.cum( mia, subset=c("sex","age","cept","onal"),
+                 ctr.mat=cbind(sx,a,1,1,CM), intl=intl, ci.Exp=TRUE )
+ tabSurv["Int",2-sx,paste(a),,] <- exp(-chI[,c(1,3,2)])
+ tabSurv["Cnv",2-sx,paste(a),,] <- exp(-chC[,c(1,3,2)])
+ pls(yaxt=if(a>47) "n" else "s",xl="")
+ if( sx==1 )
+ text( 18, 0.99, paste(a), font=1, adj=c(1,1), cex=1.5 )
+ abline(h=1:19/20,col="white")
+    }
> mtext( "Survival", side=2, line=2, las=0, outer=TRUE )
> mtext( "Time since baseline (years)", side=1, line=0, las=0, outer=TRUE )
> mtext( "Age at entry", side=3, line=0, las=0, outer=TRUE )
> mtext( "Men", at=0.75, side=2, line=2, las=0, outer=TRUE )
> mtext( "Women", at=0.25, side=2, line=2, las=0, outer=TRUE )
```

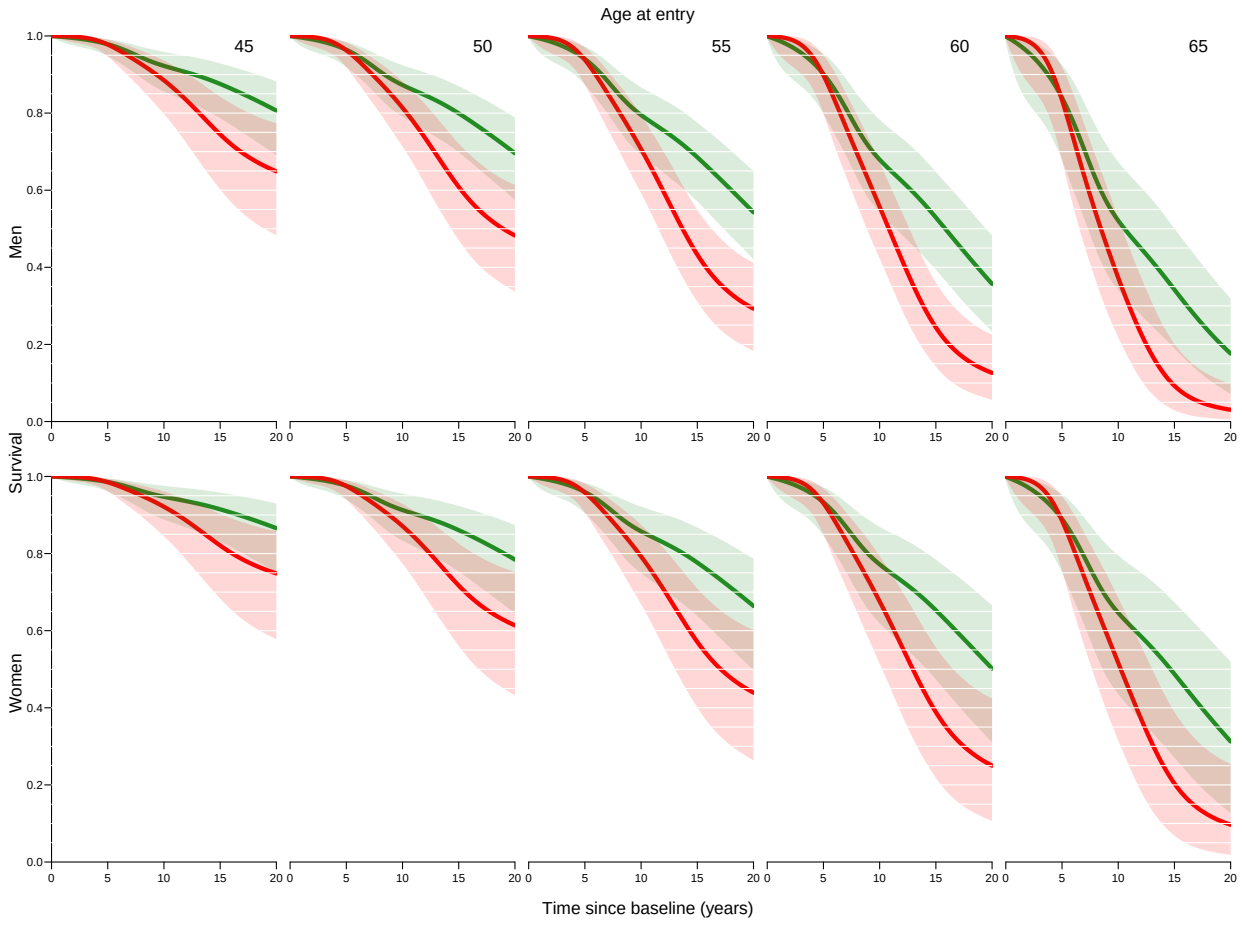

Figure 3.7: *The effect of intervention and age on survival. Estimated survival probabilities with 95% c.i. for patients entering the study in different ages.*

3.2.6 Change in mortality after intervention

Formally we have already tested whether the hazard ratio changed at the end of the intervention period — it was the test of proportionality of hazard rates between the groups,

but we could of course do a more naive test by expanding the model with proportional hazards with an interaction between allocation and an indicator of intervention period:

```
> ms <- update( m0, . ~ . + I(tsb>7.8 & allocation=="Conventional") )
> round( ci.exp( ms, subset="Conventional" ), 3 )
                                                exp(Est.)  2.5% 97.5%
allocationConventional                              1.149 0.567 2.326
I(tsb > 7.8 & allocation == "Conventional")TRUE     1.932 0.850 4.393
> zz <- ci.exp( ms, subset="Conventional",
+                   ctr.mat=rbind("0-7.8 "=c(1,0),
+                                 "  7.8+"=c(1,1),
+                                 "ratio "=c(0,1)) )
> round( cbind( zz, 1/zz[,c(1,3,2)] ), 3 )
       exp(Est.)  2.5% 97.5% exp(Est.) 97.5%  2.5%
0-7.8      1.149 0.567 2.326     0.870 0.430 1.762
  7.8+     2.220 1.366 3.606     0.450 0.277 0.732
ratio      1.932 0.850 4.393     0.518 0.228 1.177
```

So we see that if anything, the gap between intensive and conventional becomes wider after the end of intervention, although they are not formally statistically significantly different — the confidence interval for the ratio of HRs before and after is (0.85;4.39).

However, the basic assumption in this type of model is that there is a constant mortality ratio between the two groups, and that this ratio suddenly changes to a different value a exactly the time when the two groups are offered the same treatment. This basically amounts to an averaging of the HR in the two intervals defined as during (0–7.8 years) and after (7.8+ years) intervention, which is apparently what was done in the NEJM paper from 2008 [1]. However the estimated mortality curves and the HR curve as a function of time since baseline as shown in figure 3.3 does not indicate any dramatic change in mortality post intervention, although it seems that the mortality rates do converge after the intervention period.

3.2.7 CVD events and all-cause mortality

It is of course also of interest to see whether CVD events post baseline have any influence on mortality rates. However, by subdividing follow-up by CVD-event status we are conditioning on intermediate events that are also influenced by the intervention, and by that token we are likely to underestimate the effect of allocation.

So we try to add indicators of intermediate CVD to the PH model for all-cause mortality:

- `m1` effect any CVD after baseline
- `ml` additional effect of subsequent CVD events, the same effect of each extra CVD event (up to 3)
- `m3` separate effects of extra CVDs — that is no assumption of the same effect of each new CVD

```
> ( anyCVD <- levels(S1$lex.Cst)[2:4] )
[1] "1st CVD" "2nd CVD" "3rd CVD"
> m1 <- update( m0, . ~ . + I(lex.Cst %in% anyCVD) )
> ml <- update( m1, . ~ . + pmax(as.integer(lex.Cst)-2,0) )
> m3 <- update( mi, . ~ . + lex.Cst )
> round( ci.exp( m1, subset="Cst" ), 4 )
```

```
                           exp(Est.)  2.5% 97.5%
I(lex.Cst %in% anyCVD)TRUE    4.9295 3.155 7.702
> round( ci.exp( m1, subset="Cst" ), 4 )
                                  exp(Est.)   2.5%  97.5%
I(lex.Cst %in% anyCVD)TRUE           3.8140 2.3035 6.3150
pmax(as.integer(lex.Cst) - 2, 0)     1.5218 1.1190 2.0694
> round( ci.exp( m3, subset="Cst" ), 4 )
               exp(Est.)   2.5%   97.5%
lex.Cst1st CVD    3.6949 2.1884  6.2384
lex.Cst2nd CVD    6.3278 3.3742 11.8666
lex.Cst3rd CVD    7.9542 4.1698 15.1730
```

We can summarize the HR of death versus the group with no post-baseline CVD from the three models:

```
> cmpeff <-  cbind( ci.exp( m3, subset="Cst" ),
+                   ci.exp( m1, subset="Cst",
+                               ctr.mat=rbind(c(1,0),
+                                             c(1,1),
+                                             c(1,2)) ),
+                   ci.exp( m1, subset="Cst" )[c(1,1,1),] )
> colnames( cmpeff )[c(1,4,7)] <- c("Grouped","Linear","Uniform")
> round( cmpeff, 3 )
               Grouped  2.5%  97.5% Linear  2.5%  97.5% Uniform  2.5% 97.5%
lex.Cst1st CVD   3.695 2.188  6.238  3.814 2.303  6.315   4.929 3.155 7.702
lex.Cst2nd CVD   6.328 3.374 11.867  5.804 3.663  9.196   4.929 3.155 7.702
lex.Cst3rd CVD   7.954 4.170 15.173  8.832 4.854 16.070   4.929 3.155 7.702
```

We can also compare the models, first by a Wald test for linearity of effects and also by comparing different models using a likelihood ratio test:

```
> Wald( m3, subset="Cst", ctr.mat=rbind(c(1,-2,1)) )
    Chisq      d.f.         P
0.2536462 1.0000000 0.6145193
> anova( m0, m1, m1, m3, test="Chisq" )
Analysis of Deviance Table

Model 1: (lex.Xst %in% dst) ~ allocation + Ns(tsb, kn = t.kn)
Model 2: (lex.Xst %in% dst) ~ allocation + Ns(tsb, kn = t.kn) + I(lex.Cst %in%
    anyCVD)
Model 3: (lex.Xst %in% dst) ~ allocation + Ns(tsb, kn = t.kn) + I(lex.Cst %in%
    anyCVD) + pmax(as.integer(lex.Cst) - 2, 0)
Model 4: (lex.Xst %in% dst) ~ allocation + lex.Cst + allocation:Ns(tsb,
    kn = t.kn)
  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1     29243     1234.3
2     29242     1182.2  1   52.038 5.445e-13
3     29241     1175.3  1    6.884  0.008698
4     29235     1172.3  6    2.982  0.811160
```

The three p-values are for the hypotheses of no effect of CVD, linear effect of no. CVD event, and additional non-linear effect of CVD.

Thus we see that the increase in mortality is the same from 0 to 1, from 1 to 2 and from 2 to 3 CVD events after baseline, namely a HR of 1.52 (1.12;2.07) (from model `m1`), and the likelihood ratio test for linearity, $p = 0.811$.

However, it might be of more interest to see if these effects are the same in the two intervention groups:

```
> mx1 <- update( m0, . ~ . + allocation:I(lex.Cst %in% levels(lex.Cst)[2:4]) )
> mxl <- update( mx1,. ~ . + allocation:pmax(as.integer(lex.Cst)-2,0) )
> mx3 <- update( m0, . ~ . + allocation:lex.Cst )
> round( ci.exp( ml ), 4 )
                                exp(Est.)   2.5%     97.5%
(Intercept)                        0.0026 0.0004    0.0192
allocationConventional             1.1395 0.7360    1.7640
Ns(tsb, kn = t.kn)1               10.3812 1.5536   69.3676
Ns(tsb, kn = t.kn)2                7.8597 0.7476   82.6293
Ns(tsb, kn = t.kn)3                6.5515 1.3786   31.1337
Ns(tsb, kn = t.kn)4               40.7542 0.5228 3176.6372
Ns(tsb, kn = t.kn)5                2.0692 0.9116    4.6968
I(lex.Cst %in% anyCVD)TRUE         3.8140 2.3035    6.3150
pmax(as.integer(lex.Cst) - 2, 0)   1.5218 1.1190    2.0694
> round( ci.exp( mx1, subset=c("base","Cst") ), 4 )
                                                                 exp(Est.)   2.5%   97.5%
allocationIntensive:I(lex.Cst %in% levels(lex.Cst)[2:4])TRUE        5.8950 3.0955 11.2265
allocationConventional:I(lex.Cst %in% levels(lex.Cst)[2:4])TRUE     4.2163 2.3348  7.6141
> round( ci.exp( mxl, subset=c("base","Cst") ), 4 )
                                                                 exp(Est.)   2.5%   97.5%
allocationIntensive:I(lex.Cst %in% levels(lex.Cst)[2:4])TRUE        5.5162 2.7060 11.2451
allocationConventional:I(lex.Cst %in% levels(lex.Cst)[2:4])TRUE     2.8012 1.4066  5.5786
allocationIntensive:pmax(as.integer(lex.Cst) - 2, 0)                1.1895 0.6485  2.1819
allocationConventional:pmax(as.integer(lex.Cst) - 2, 0)             1.7054 1.1842  2.4560
> round( ci.exp( mx3, subset=c("base","Cst") ), 4 )
                                       exp(Est.)   2.5%   97.5%
allocationIntensive:lex.Cst1st CVD        5.5537 2.6919 11.4578
allocationConventional:lex.Cst1st CVD     2.6120 1.2672  5.3841
allocationIntensive:lex.Cst2nd CVD        6.2803 2.2685 17.3868
allocationConventional:lex.Cst2nd CVD     5.8345 2.6752 12.7245
allocationIntensive:lex.Cst3rd CVD        7.8775 2.2323 27.7987
allocationConventional:lex.Cst3rd CVD     7.5113 3.5537 15.8763
> anova( m1, mx1, test="Chisq" )[2,-(1:2),drop=F]
  Df Deviance Pr(>Chi)
2  1  0.57058     0.45
> anova( ml, mxl, test="Chisq" )[2,-(1:2),drop=F]
  Df Deviance Pr(>Chi)
2  2   1.9986   0.3681
> anova( m3, mx3, test="Chisq" )[2,-(1:2),drop=F]
  Df Deviance Pr(>Chi)
2 -2 -0.43183   0.8058
```

. . . and indeed they are; so the relevant effect is still the overall hazard ratio from a proportional hazards model that takes these different effects into account; here we show the allocation effects both ways as well as the effects of post-baseline CVD:

```
> msb <- update( m0 , . ~ . + I(age-tsb) + sex )
> msl <- update( msb, . ~ . + I(lex.Cst %in% levels(lex.Cst)[2:4])
+                           + pmax(as.integer(lex.Cst)-2,0) )
> msf <- update( msb, . ~ . + lex.Cst )
> zz <- ci.exp( msb, subset=c("allo","Cst") )
> round( rbind( 1/zz[1,c(1,3,2)], zz ), 3 )
                     exp(Est.) 97.5%  2.5%
                         0.526 0.346 0.800
allocationConventional   1.899 1.249 2.888
> zz <- ci.exp( msl, subset=c("allo","Cst") )
> round( rbind( 1/zz[1,c(1,3,2)], zz ), 3 )
                                       exp(Est.) 97.5%  2.5%
                                           0.830 0.533 1.293
allocationConventional                     1.205 0.774 1.877
I(lex.Cst %in% levels(lex.Cst)[2:4])TRUE   3.003 1.799 5.012
pmax(as.integer(lex.Cst) - 2, 0)           1.577 1.151 2.160
```

```
> zz <- ci.exp( msf, subset=c("allo","Cst") )
> round( rbind( 1/zz[1,c(1,3,2)], zz ), 3 )
                        exp(Est.) 97.5%   2.5%
                            0.832 0.534  1.297
allocationConventional      1.201 0.771  1.872
lex.Cst1st CVD              3.063 1.814  5.172
lex.Cst2nd CVD              4.422 2.360  8.286
lex.Cst3rd CVD              7.743 4.103 14.612
```

We see that as soon as post-baseline CVD is included in the model in one form or other, the effect of allocation group is changed from a HR of 1.9 to about 1.2, the latter non-significant. This is because the major part of the intervention effect is exercised on the transitions from DM to CVD and to subsequent CVD events, and these transitions are not modelled in the overall mortality analysis.

We shall return to a more detailed analysis of the CVD event rates later.

3.2.8 Summary of all-cause mortality

For all-cause mortality there is no indication of non-proportional hazards over the follow-up range (P=0.447), and the hazard ratio between the conventional and the intensive group is 1.94 (1.28;2.96).

Persons who see a post-baseline CVD have a higher mortality (HR 3.00 (1.80;5.58)) and an increase in HR of 1.54 (1.12;2.10) per extra CVD event up to 3; these effects are the same for the two intervention groups too.

The M/F HR was 1.49 (0.90;2.48) and effect of age was a HR of 1.11 (1.07;1.15) per year, corresponding to a doubling time of mortality of 6.8 (5.9;10.9) years. The latter shows that summary survival figures are heavily influenced by the particular age-distribution in the study population, *i.e.* they cannot easily be generalized to other settings, because their very strong dependency on the particular age and sex-distribution in the Steno 2 study. Hence the reporting of survival by age and sex.

The 20-year survival in the intensive group is 55%, and in the standard group 32%. These figures are however specific for the particular age-composition in the Steno 2 patient population. It is more relevant to state the 20-year survival of for example men entering at age 50 as 72% and 52% respectively; a more complete overview of survival probabilities is here:

```
> ftable( round(tabSurv[,,,c("5","10"),]*100,1),
+         col.vars=c(4,2,5),
+         row.vars=c(3,1) )
        tfr     5                           10
        sex     M              F             M              F
        what  est   lo   up  est   lo   up  est   lo   up  est   lo   up
age grp
45  Int      97.9 94.8 99.1 98.6 96.1 99.5 92.2 85.6 95.9 94.7 88.9 97.6
    Cnv      97.8 94.7 99.1 98.5 96.2 99.4 88.5 79.8 93.6 92.1 84.5 96.1
50  Int      96.4 91.9 98.4 97.6 94.0 99.0 87.3 79.1 92.4 91.3 83.4 95.5
    Cnv      96.3 91.9 98.3 97.5 94.1 99.0 81.4 71.3 88.2 87.1 77.4 92.8
55  Int      94.0 87.3 97.2 96.0 90.4 98.3 79.5 69.2 86.7 85.8 74.9 92.2
    Cnv      93.8 87.3 97.0 95.8 90.6 98.2 70.7 59.2 79.4 79.2 66.8 87.4
60  Int      90.2 79.8 95.3 93.3 84.6 97.2 67.9 54.3 78.3 77.2 61.9 86.9
    Cnv      89.8 79.8 95.0 93.0 85.0 96.9 55.7 42.2 67.2 67.6 51.5 79.4
65  Int      84.0 67.8 92.5 89.0 75.0 95.3 52.1 34.1 67.4 64.6 43.5 79.5
    Cnv      83.4 67.8 91.9 88.6 75.6 94.9 37.3 21.9 52.7 51.7 31.5 68.5
> ftable( round(tabSurv[,,,c("15","20"),]*100,1),
+         col.vars=c(4,2,5),
+         row.vars=c(3,1) )
```

```
        tfr    15                          20
        sex     M             F             M             F
        what  est   lo   up  est   lo   up  est   lo   up  est   lo   up
age grp
45  Int      87.5 78.7 92.9 91.5 83.1 95.8 80.6 69.0 88.2 86.6 75.1 93.0
    Cnv      74.5 60.2 84.3 82.1 68.3 90.3 64.9 48.3 77.3 74.9 57.8 85.8
50  Int      79.9 69.8 87.0 86.1 75.3 92.4 69.6 57.4 78.9 78.4 64.5 87.4
    Cnv      60.9 47.0 72.1 71.7 56.0 82.6 48.2 33.7 61.4 61.4 43.3 75.3
55  Int      68.5 57.1 77.5 77.7 63.8 86.7 54.3 42.0 65.0 66.4 49.7 78.7
    Cnv      43.3 31.1 54.9 57.1 40.1 70.9 29.3 18.3 41.2 43.9 26.4 60.2
60  Int      52.9 39.4 64.7 65.3 47.4 78.4 35.7 23.3 48.3 50.2 31.0 66.6
    Cnv      24.4 14.2 36.0 38.9 21.8 55.7 12.6  5.6 22.6 25.0 10.6 42.4
65  Int      34.2 19.0 50.0 48.8 27.1 67.3 17.6  7.1 32.0 31.3 12.7 52.0
    Cnv       9.3  2.9 20.1 20.4  6.8 38.9  3.1  0.5  9.9  9.7  1.9 25.4
```

3.3 CVD and other cause mortality

Unlike the analysis of overall survival there is not a simple transformation of rates that yields the survival function when we consider multiple causes of death (CVD / non-CVD). Moreover, the survival function might not be the target of interest when analyzing CVD and other mortality separately — more likely the probabilities of interest would be the fraction of patients that die from each cause; or possibly the fraction of patients that see 0, 1, 2 or 3 CVD events before death. We shall return to these in the next chapter, as they depend on full modeling of all transitions shown in figure 1.3.

The analysis of the cause-specific mortalities is in principle exactly parallel to that of all-cause mortality, except that we use two different subsets of the death states as events:

```
> ( dst <- levels( S1 )[-(1:4)] )
[1] "Dead 0" "CV-D 0" "Dead 1" "CV-D 1" "Dead 2" "CV-D 2" "Dead 3" "CV-D 3"
```

3.3.1 CVD mortality

For the CVD mortality, the events are every other of the states:

```
> dst
[1] "Dead 0" "CV-D 0" "Dead 1" "CV-D 1" "Dead 2" "CV-D 2" "Dead 3" "CV-D 3"
> ( dcv <- dst[1:4*2] )
[1] "CV-D 0" "CV-D 1" "CV-D 2" "CV-D 3"
> sum( S1$lex.Xst %in% dcv )
[1] 38
```

and this is used in the modeling, but note that we only have 38 events, so the modeling possibilities are smaller, and we therefore only use 4 knots for the hazard rate function:

```
> ( t.kn <- with( subset(S1,lex.Xst %in% dcv),
+                 c(0,quantile( tsb+lex.dur, probs=(1:3-0.5)/3) )) )
         16.66667%       50% 83.33333%
 0.000000  4.935889  9.639973 16.059776
> m0 <- glm( (lex.Xst %in% dcv) ~ allocation + Ns(tsb, kn=t.kn),
+            family = poisson,
+            offset = log(lex.dur),
+              data = S1 )
> mi <- update( m0, . ~ . - Ns(tsb, kn=t.kn) +
+                 allocation:Ns(tsb, kn=t.kn) )
> ma  <- update( m0, . ~ . + I(age-tsb) + sex )
> mia <- update( mi, . ~ . + I(age-tsb) + sex )
> anova( ma, m0, mi, mia, ma, test="Chisq" )[-1,-(1:2)]
```

```
  Df Deviance  Pr(>Chi)
2 -2 -23.7778 6.866e-06
3  3   8.2851   0.04047
4  2  24.3871 5.063e-06
5 -3  -8.8944   0.03073
```

```
> HR.IC <- rbind( ci.exp( m0, subset="allo", pval=TRUE ),
+                 ci.exp( ma, subset="allo", pval=TRUE ) )
> round( cbind( 1/HR.IC[,c(1,3,2)], HR.IC ), 3 )
```

```
                       exp(Est.) 97.5%  2.5% exp(Est.)  2.5% 97.5%     P
allocationConventional     0.379 0.191 0.754     2.637 1.327 5.240 0.006
allocationConventional     0.353 0.177 0.705     2.832 1.419 5.652 0.003
```

```
> round( cbind( ci.exp( ma , subset=c("age","sex") ),
+               ci.exp( mia, subset=c("age","sex") ) ), 3 )
```

```
             exp(Est.)  2.5% 97.5% exp(Est.)  2.5% 97.5%
I(age - tsb)     1.123 1.064 1.186     1.126 1.066 1.189
sexM             2.256 0.940 5.415     2.266 0.944 5.442
```

We see that there is a borderline significant interaction, so formally we can not maintain the CVD mortality rates are proportional throughout the study period. Ignoring this, the estimated overall HR between intensive and conventional is 0.38 (0.19;0.75); and even with the rather small number of events this is significantly different from 1.

We then update the HR collector

```
> dimnames( mainCI )[[1]][2]
[1] "CVD mortality"
```

```
> zz <- HR.IC
> zz[,1:3] <- 1/HR.IC[,c(1,3,2)]
> mainCI[2,,] <- zz
> save( mainCI, file="../data/mainCI.Rda" )
```

It is of course also of interest to see whether CVD event post baseline only influence CVD mortality rates; so we try to add these to the model:

```
> mx <- update( mia, . ~ . + I(lex.Cst %in% levels(lex.Cst)[2:4]) )
> round( ci.exp( mx ), 3 )
```

```
                                           exp(Est.)  2.5%    97.5%
(Intercept)                                    0.000 0.000    0.001
allocationConventional                         0.402 0.010   16.487
I(age - tsb)                                   1.110 1.045    1.178
sexM                                           2.118 0.879    5.104
I(lex.Cst %in% levels(lex.Cst)[2:4])TRUE       7.182 3.191   16.164
allocationIntensive:Ns(tsb, kn = t.kn)1        0.238 0.024    2.371
allocationConventional:Ns(tsb, kn = t.kn)1     8.159 1.121   59.412
allocationIntensive:Ns(tsb, kn = t.kn)2        1.242 0.008  193.489
allocationConventional:Ns(tsb, kn = t.kn)2    12.365 0.024 6368.653
allocationIntensive:Ns(tsb, kn = t.kn)3        0.438 0.102    1.875
allocationConventional:Ns(tsb, kn = t.kn)3     4.002 0.988   16.207
```

```
> anova( mia, mx, test="Chisq" )
```

```
Analysis of Deviance Table

Model 1: (lex.Xst %in% dcv) ~ allocation + I(age - tsb) + sex + allocation:Ns(tsb,
    kn = t.kn)
Model 2: (lex.Xst %in% dcv) ~ allocation + I(age - tsb) + sex + I(lex.Cst %in%
    levels(lex.Cst)[2:4]) + allocation:Ns(tsb, kn = t.kn)
  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1     29240     555.60
2     29239     527.83  1   27.771 1.365e-07
```

We see that there is a significant increased CVD mortality in persons with CVD events post baseline.

We now plot the two mortality curves and the RR between them from the models where follow-up CVD events are not taken into account; for the sex-age controlled model we use men aged 55 to show the mortality rates (median age at baseline is 56.6 years):

```
> CM <- Ns( t.pt, kn=t.kn )
> nd <- data.frame(tsb=t.pt,lex.dur=1000,age=55+t.pt,sex="M")
> rS.t <- ci.pred( mi, newdata=cbind(nd,allocation="Conventional") )
> rI.t <- ci.pred( mi, newdata=cbind(nd,allocation="Intensive") )
> rS.a <- ci.pred( mia, newdata=cbind(nd,allocation="Conventional") )
> rI.a <- ci.pred( mia, newdata=cbind(nd,allocation="Intensive") )
> pS.t <- ci.pred( m0, newdata=cbind(nd,allocation="Conventional") )
> pI.t <- ci.pred( m0, newdata=cbind(nd,allocation="Intensive") )
> pS.a <- ci.pred( ma, newdata=cbind(nd,allocation="Conventional") )
> pI.a <- ci.pred( ma, newdata=cbind(nd,allocation="Intensive") )
> # Check parameter sequencing using subset=
>  round( ci.exp( mia, subset=c("Intensive","Conventional") ), 3 )
                                          exp(Est.)  2.5%     97.5%
allocationIntensive:Ns(tsb, kn = t.kn)1       0.388 0.039     3.865
allocationIntensive:Ns(tsb, kn = t.kn)2       5.676 0.043   757.559
allocationIntensive:Ns(tsb, kn = t.kn)3       0.661 0.158     2.764
allocationConventional                        0.524 0.014    19.804
allocationConventional:Ns(tsb, kn = t.kn)1    20.276 2.918   140.910
allocationConventional:Ns(tsb, kn = t.kn)2    81.812 0.198 33882.038
allocationConventional:Ns(tsb, kn = t.kn)3     7.729 1.983    30.122
> RR.t <- ci.exp( mi , subset=c("Intensive","Conventional"), ctr.mat=cbind(CM,-1,-CM) )
> RR.a <- ci.exp( mia, subset=c("Intensive","Conventional"), ctr.mat=cbind(CM,-1,-CM) )
```

In order to plot the rightmost panel for the ESM together with the same for the all cause mortality we save the relevant structures:

```
> save( t.pt, HR.IC,
+             rI.t, rS.t,
+             rI.a, rS.a,
+             pI.t, pS.t, RR.t,
+             pI.a, pS.a, RR.a,
+       file = "../data/cvd-mort.Rda" )
```

We now have the predicted mortality rates and the estimated HRs from the two models, so we can plot them side by side and see to what extent adjustment for sex and age at baseline influences the estimates:

```
> par( mfrow=c(1,2), mar=c(0,0,0,0), oma=c(3,4,1,2), mgp=c(3,1,0)/1.6 )
> yl <- c(0.1,250)
> matplot( t.pt, cbind( rI.t, rS.t, RR.t),
+          log="y", ylim=yl, xlab="",
+          type="l", lty=1, lwd=c(3,1,1), col=rep(clr,each=3) )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[1,-4], col="gray" )
> abline( h=1/HR.IC[1,1], col="gray", lwd=3 )
> matplot( t.pt, cbind(rI.a,rS.a,RR.a),
+          log="y", ylim=yl, yaxt="n", ylab="",
+          type="l", lty=1, lwd=c(3,1,1), col=rep(clr,each=3) )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[2,-4], col="gray" )
> abline( h=1/HR.IC[2,1], col="gray", lwd=3 )
> mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
> text( rep(9,3), 20*0.75^c(2,1,3), c("Intensive","Conventional","Rate ratio"),
+       col=clr, font=2, cex=1.2, adj=0 )
> mtext( "CVD mortality rate per 1000 PY / Hazard ratio",
+        side=2, line=3, las=0, outer=TRUE )
```

There is a decrease in the intensive vs. conventional HR of CVD mortality from around 5 years since randomization, consonant with the significant non-proportionality (p=0.040).

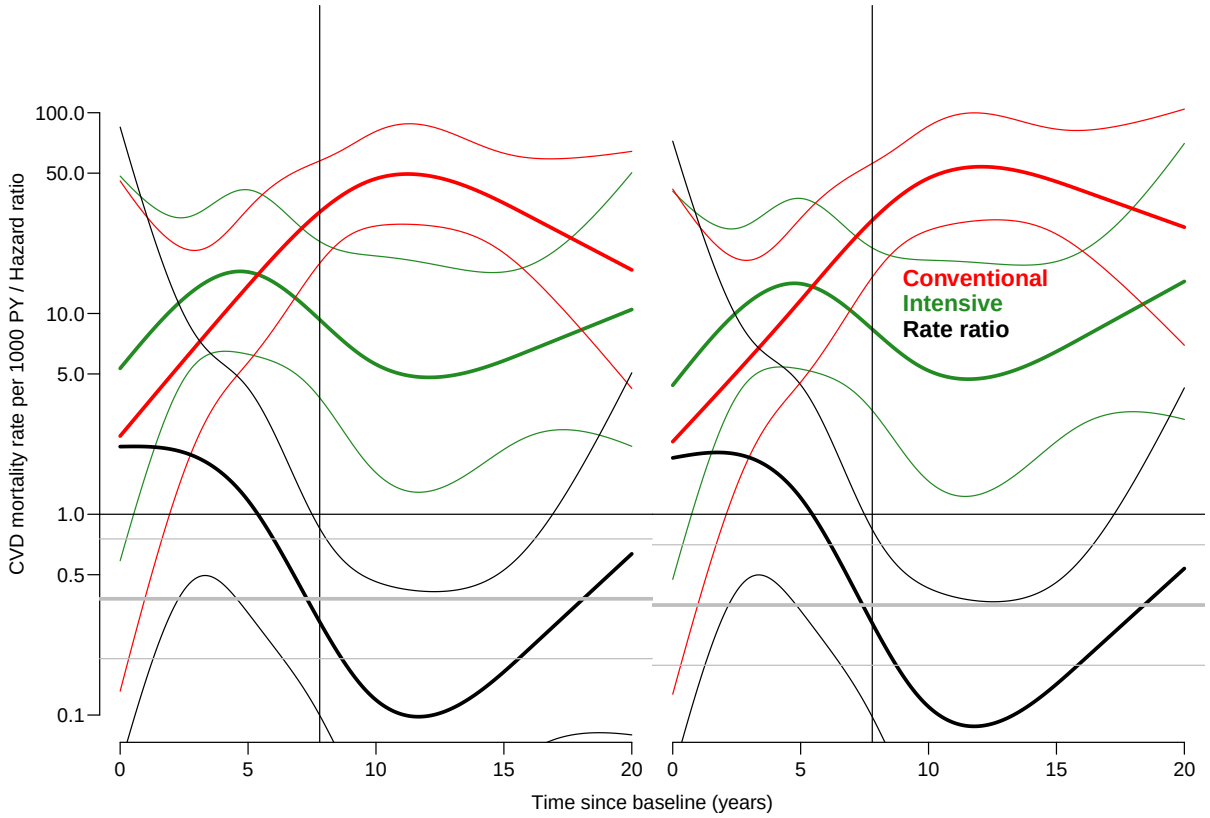

Figure 3.8: *CVD mortality rates and hazard ratio between the two groups. The gray lines indicate the overall HR with 95% c.i., the vertical line indicates the intervention end. Left panel is a model with only group and time since baseline, the model in the right panel also includes sex and age at baseline as a linear term on the log-mortality scale, and mortality rates are shown for a man aged 55 at baseline. Horizontal gray lines indicate the overall HR from the (non-tenable) proportional hazads models.*

3.3.2 Non-CVD mortality

For the non-CVD mortality, the events are also every other of the states:

```
> ( doth <- dst[0:3*2+1] )
[1] "Dead 0" "Dead 1" "Dead 2" "Dead 3"
> sum( S1$lex.Xst %in% doth )
[1] 55
```

and this is used in the modeling, but note that we only have 55 events, so the modeling possibilities are smaller than for all-cause mortality:

```
> ( t.kn <- with( subset(S1,lex.Xst %in% doth),
+                 c(0,quantile( tsb+lex.dur, probs=(1:4-0.5)/4) )) )
             12.5%     37.5%     62.5%     87.5%
 0.000000  5.535250  9.140999 12.841889 17.458590
> m0 <- glm( (lex.Xst %in% doth) ~ allocation + Ns(tsb, kn=t.kn),
+            family = poisson,
+            offset = log(lex.dur),
+              data = S1 )
> mi <- update( m0, . ~ . - Ns(tsb, kn=t.kn) +
+                 allocation:Ns(tsb, kn=t.kn) )
> ma  <- update( m0, . ~ . + I(age-tsb) + sex )
> mia <- update( mi, . ~ . + I(age-tsb) + sex )
> anova( ma, m0, mi, mia, ma, test="Chisq" )[-1,-(1:2)]
```

```
  Df Deviance  Pr(>Chi)
2 -2 -21.0909 2.631e-05
3  4   1.7994    0.7726
4  2  20.9958 2.759e-05
5 -4  -1.7043    0.7899
```

```
> HR.IC <- rbind( ci.exp( m0, subset="allo", pval=TRUE ),
+                 ci.exp( ma, subset="allo", pval=TRUE ) )
> rownames(HR.IC)<- c("Allo|ph","Allo|ph+age+sex")
> round( cbind( HR.IC, 1/HR.IC[,c(1,3,2)] ), 3 )
```

```
                exp(Est.)  2.5% 97.5%     P exp(Est.) 97.5%  2.5%
Allo|ph             1.422 0.835 2.421 0.195     0.703 0.413 1.197
Allo|ph+age+sex     1.459 0.853 2.494 0.168     0.686 0.401 1.172
```

```
> round( cbind( ci.exp( ma , subset=c("age","sex") ),
+               ci.exp( mia, subset=c("age","sex") ) ), 3 )
```

```
             exp(Est.)  2.5% 97.5% exp(Est.)  2.5% 97.5%
I(age - tsb)     1.100 1.054 1.149     1.100 1.054 1.149
sexM             1.149 0.624 2.114     1.164 0.632 2.143
```

We see that there is no significant interaction, so formally we can maintain the the mortality rates are proportional throughout the study period, and the estimated HR between intensive and conventional is 0.70 (0.41;1.19) — not significantly different from 1.

We then update the HR collector

```
> dimnames( mainCI )[[1]][3]
[1] "non-CVD mortality"
```

```
> zz <- HR.IC
> zz[,1:3] <- 1/HR.IC[,c(1,3,2)]
> mainCI[3,,] <- zz
> save( mainCI, file="../data/mainCI.Rda" )
```

It is also of some interest to see whether CVD events post baseline also influences non-CVD mortality rates; so we try to add these to the model:

```
> mx <- update( mia, . ~ . + I(lex.Cst %in% levels(lex.Cst)[2:4]) )
> anova( ma, mx, test="Chisq" )[-1,-(1:2)]
  Df Deviance Pr(>Chi)
2  5   15.288   0.0092
```

```
> round( ci.exp( mx, subset="Cst" ), 3 )
```

```
                                          exp(Est.)  2.5% 97.5%
I(lex.Cst %in% levels(lex.Cst)[2:4])TRUE      2.878 1.638 5.055
```

We see that there is a significant effect of CVD events post baseline; any CVD event holds a HR of 2.9 (1.6;5.1).

We now plot the two mortality curves and the HR between them from the models where post follow-up CVD events are not taken into account; for the sex-age-model we use men aged 55 to show the mortality rates (median age at baseline is 56.6 years):

```
> CM <- Ns( t.pt, kn=t.kn )
> nd <- data.frame(tsb=t.pt,lex.dur=1000,age=55+t.pt,sex="M")
> rS.t <- ci.pred( mi , newdata=cbind(nd,allocation="Conventional") )
> rI.t <- ci.pred( mi , newdata=cbind(nd,allocation="Intensive") )
> rS.a <- ci.pred( mia, newdata=cbind(nd,allocation="Conventional") )
> rI.a <- ci.pred( mia, newdata=cbind(nd,allocation="Intensive") )
> # Check parameter sequencing using subset=
>  round( ci.exp( mi , subset=c("Intensive","Conventional") ), 3 )
```

```
                                          exp(Est.)  2.5%        97.5%
allocationIntensive:Ns(tsb, kn = t.kn)1      254.255 0.544 1.188090e+05
allocationIntensive:Ns(tsb, kn = t.kn)2       23.663 0.290 1.930249e+03
allocationIntensive:Ns(tsb, kn = t.kn)3     6575.244 0.007 6.102874e+09
allocationIntensive:Ns(tsb, kn = t.kn)4       12.715 1.173 1.378550e+02
allocationConventional                         0.922 0.000 3.639632e+03
allocationConventional:Ns(tsb, kn = t.kn)1   153.879 1.403 1.687202e+04
allocationConventional:Ns(tsb, kn = t.kn)2    50.276 1.523 1.660002e+03
allocationConventional:Ns(tsb, kn = t.kn)3 50603.940 0.964 2.655259e+09
allocationConventional:Ns(tsb, kn = t.kn)4     6.731 1.058 4.281300e+01
```

```
> RR.t <- ci.exp( mi , subset=c("Intensive","Conventional"), ctr.mat=cbind(CM,-1,-CM) )
> RR.a <- ci.exp( mia, subset=c("Intensive","Conventional"), ctr.mat=cbind(CM,-1,-CM) )
```

We now have the predicted mortality rates and the estimated HRs from the two models, so we can plot them side by side and see to what extent adjustment for sex and age at baseline influences the estimates:

```
> par( mfrow=c(1,2), mar=c(0,0,0,0), oma=c(3,4,1,2), mgp=c(3,1,0)/1.6 )
> yl <- c(0.1,250)
> matplot( t.pt, cbind(rI.t,rS.t,RR.t),
+          log="y", ylim=yl, xlab="",
+          type="l", lty=1, lwd=c(3,1,1), col=rep(clr,each=3) )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[1,-4], col="gray" )
> abline( h=1/HR.IC[1,1], col="gray", lwd=3 )
> matplot( t.pt, cbind(rI.a,rS.a,RR.a),
+          log="y", ylim=yl, yaxt="n", ylab="",
+          type="l", lty=1, lwd=c(3,1,1), col=rep(clr,each=3) )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[2,-4], col="gray" )
> abline( h=1/HR.IC[2,1], col="gray", lwd=3 )
> mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
> text( rep(11,3), 10*0.75^c(2,1,3), c("Intensive","Conventional","Rate ratio"),
+       col=clr, font=2, cex=1.2, adj=0 )
> mtext( "Non-CVD mortality rate per 1000 PY / Hazard ratio",
+        side=2, line=3, las=0, outer=TRUE )
```

We see from figure 3.9 that mortality increases substantially the first 5 years, and thereafter levels off; also it is pretty obvious that there is not much evidence of non-proportionality, the p-value for the null of proportional hazards along the time scale is 0.773.

3.4 Mortality summary

In terms of mortality the majority of the intervention effect is on CVD mortality; there is no statistically or clinically (figure **??**) significant difference between the non-CVD mortality between intensive and conventional group.

The intensive effects is manifest as a crude HR of 2.6 (1.3;5.2) in the conventional group relative to the intensive group; if controlled for age and sex it is 2.8 (1.4;5.7). The HR in the intensive group relative to the conventional group is 0.37 (0.19,0.85) and 0.35 (0.18;0.70) respectively; significantly different from 1 but not very well determined.

3.4.1 Plot for the ESM

Here we just take the two summary plots for the ESM:

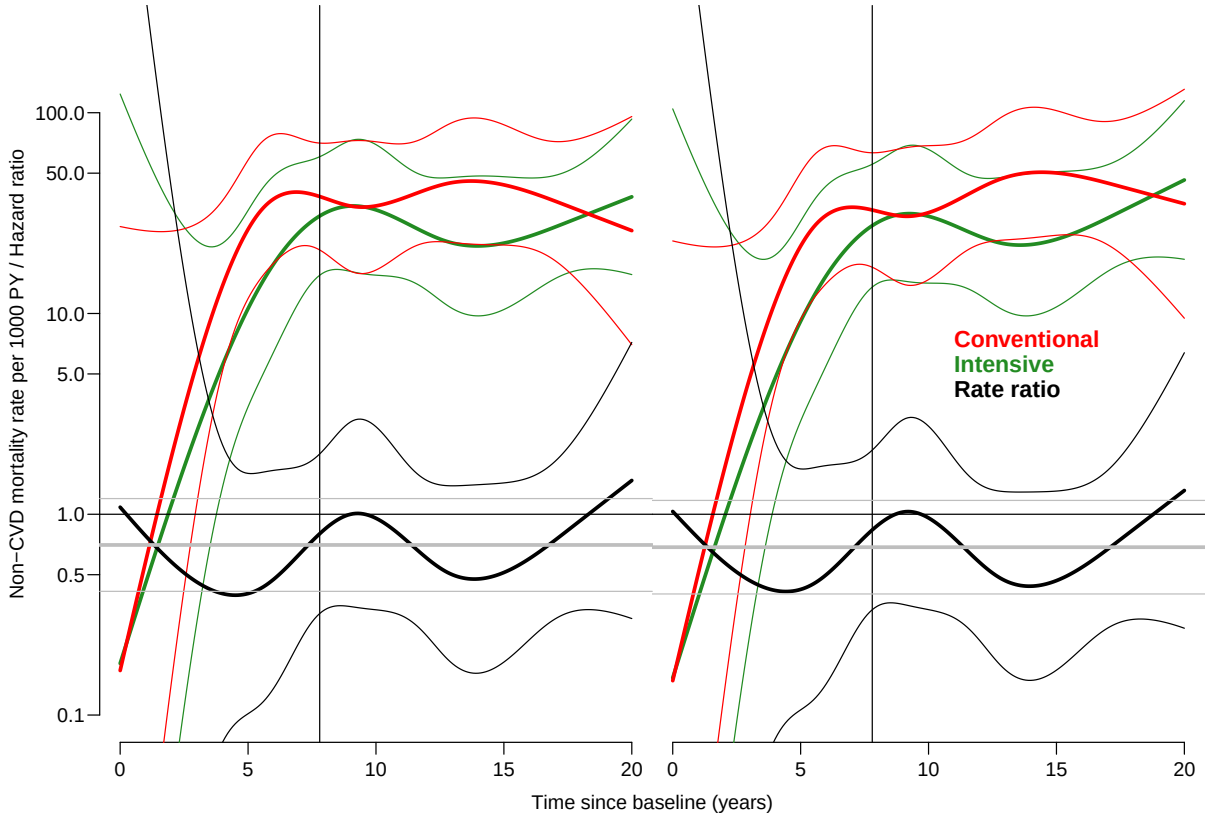

Figure 3.9: *Non-CVD mortality rates and hazard ratio between the two groups. The gray lines indicate the overall HR with 95% c.i., the vertical line indicates the intervention end. Left panel is a model with only group and time since baseline, the model in the right panel also includes sex and age at baseline as a linear term on the log-mortality scale, and mortality rates are shown for a man aged 55 at baseline. The rate-scale is the same as that for CVD mortality, but not for all-cause mortality.*

```
> clr <- c("forestgreen","red","black",gray(0.6))
> slr <- rgb( t( col2rgb(clr[1:2]) ), alpha=40, max=255 )
> par( mfrow=c(1,2), mar=c(0,2,0,0), oma=c(3,2,1,2), mgp=c(3,1,0)/1.6 )
> yl <- c(0.1,250)
> load( file = "../data/all-mort.Rda" )
> matplot( t.pt, cbind(rI.a,rS.a,RR.a),#pI.a,pS.a),
+          log="y", ylim=yl, ylab="", # yaxt="n",
+          type="l", lty=rep(c("21","63","63"),3), lend=1,
+          lwd=c(3,1,1), col=rep(clr[1:3],each=3) )
> matlines( t.pt, cbind(pI.a,pS.a),
+          type="l", lty=1, lwd=c(3,1,1), col=rep(clr[1:2],each=3) )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[2,2:3], col=clr[3], lty=1 )
> abline( h=1/HR.IC[2,  1], col=clr[3], lwd=3, lty=1 )
> mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
> text( rep(11,3), 10*0.70^c(2,1,3), c("Intensive","Conventional","Rate ratio"),
+       col=clr[-4], font=2, cex=1.2, adj=0 )
> text( 0, 200, "All cause\nmortality", adj=0, font=2, cex=1.2 )
> mtext( "Mortality rate per 1000 PY / Hazard ratio Intensive vs Conventional",
+        side=2, line=2.5, las=0, outer=FALSE )
> load( file = "../data/cvd-mort.Rda" )
> matplot( t.pt, cbind(rI.a,rS.a,RR.a),
+          log="y", ylim=yl, ylab="", yaxt="n",
+          type="l", lty=rep(c("21","63","63"),3), lend=1,
```

```
+          lwd=c(3,1,1), col=rep(clr[1:3],each=3) )
> matlines( t.pt, cbind(pI.a,pS.a),
+          type="l", lty=1, lwd=c(3,1,1), col=rep(clr[1:2],each=3) )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[2,2:3], col=clr[3], lty=1 )
> abline( h=1/HR.IC[2,  1], col=clr[3], lwd=3, lty=1 )
> mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
> text( 0, 200, "CVD\nmortality", adj=0, font=2, cex=1.2 )


> clr <- c("forestgreen","red","black",gray(0.6))
> slr <- rgb( t( col2rgb(clr[1:2]) ), alpha=40, max=255 )
> par( mfrow=c(1,2), mar=c(0,2,0,0), oma=c(3,2,1,2), mgp=c(3,1,0)/1.6 )
> yl <- c(0.1,250)
> rmcol <- c(clr[1:3],"transparent")[c(1,4,4,2,4,4,3,4,4)]
> load( file = "../data/all-mort.Rda" )
> matplot( t.pt, cbind(rI.a,rS.a,RR.a),#pI.a,pS.a),
+          log="y", ylim=yl, ylab="", # yaxt="n",
+          type="l", lty=rep(c("21","63","63"),3), lend=1,
+          lwd=c(3,1,1), col=rmcol )
> matlines( t.pt, cbind(pI.a,pS.a),
+          type="l", lty=1, lwd=c(3,1,1), col=rmcol[1:6] )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[2,  1], col=clr[3], lwd=3, lty=1 )
> mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
> text( rep(11,3), 10*0.70^c(2,1,3), c("Intensive","Conventional","Rate ratio"),
+       col=clr[-4], font=2, cex=1.2, adj=0 )
> text( 0, 200, "All cause\nmortality", adj=0, font=2, cex=1.2 )
> mtext( "Mortality rate per 1000 PY / Hazard ratio Intensive vs Conventional",
+        side=2, line=2.5, las=0, outer=FALSE )
> load( file = "../data/cvd-mort.Rda" )
> matplot( t.pt, cbind(rI.a,rS.a,RR.a),#pI.a,pS.a),
+          log="y", ylim=yl, ylab="", yaxt="n",
+          type="l", lty=rep(c("21","63","63"),3), lend=1,
+          lwd=c(3,1,1), col=rmcol )
> matlines( t.pt, cbind(pI.a,pS.a),
+          type="l", lty=1, lwd=c(3,1,1), col=rmcol[1:6] )
> abline( h=1, v=7.8 )
> abline( h=1/HR.IC[2,  1], col=clr[3], lwd=3, lty=1 )
> text( 0, 200, "CVD\nmortality", adj=0, font=2, cex=1.2 )
> mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
```

Chapter 4

CVD status during follow-up

Since a large part of the intervention effect is mediated through the occurrence of CVD events, a natural way of showing the intervention effect is through the burden of CVD events in the two groups of patients. That is, using the status of the patients during follow-up as having 0, 1, 2 or 3 CVD events, and enumerating the time spent in each state and the fraction of patients dying with 0, 1, 2 and 3 CVD events.

4.1 Multistate set-up

To this end we take the previous analysis dataset but merge the two types of death:

```
> library( survival )
> library( Epi )
> clear()
> sessionInfo()
R version 3.3.0 (2016-05-03)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 14.04.4 LTS

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C               LC_TIME=en_US.UTF-8
 [4] LC_COLLATE=en_US.UTF-8     LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8       LC_NAME=C                  LC_ADDRESS=C
[10] LC_TELEPHONE=C             LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

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

other attached packages:
[1] Epi_2.5         survival_2.39-4

loaded via a namespace (and not attached):
 [1] cmprsk_2.2-7      MASS_7.3-44       plyr_1.8.3        Matrix_1.2-1
 [5] parallel_3.3.0    etm_0.6-2         Rcpp_0.11.6       splines_3.3.0
 [9] grid_3.3.0        numDeriv_2014.2-1 lattice_0.20-31
> load( file="../data/Lx.Rda" )
> lls()
  name mode class            size
1 L0   list Lexis data.frame 160 22
2 L1   list Lexis data.frame 318 22
> levels( L0 )
[1] "DM"   "Dead" "CV-D"
> levels( L1 )
```

```
 [1] "DM"      "1st CVD" "2nd CVD" "3rd CVD" "Dead 0"  "CV-D 0"  "Dead 1"  "CV-D 1"
 [9] "Dead 2"  "CV-D 2"  "Dead 3"  "CV-D 3"
```

```
> L1 <- Relevel( L1, list("3+ CVD" = 4,
+                         "D(no CVD)" = 5:6,
+                          "D(1 CVD)" = 7:8,
+                          "D(2 CVD)" = 9:10,
+                         "D(3+ CVD)" = 11:12), first=FALSE )
```

```
      type     old       new
1  lex.Cst      DM        DM
2  lex.Cst 1st CVD   1st CVD
3  lex.Cst 2nd CVD   2nd CVD
4  lex.Cst 3rd CVD    3+ CVD
5  lex.Cst  Dead 0
6  lex.Cst  CV-D 0
7  lex.Cst  Dead 1
8  lex.Cst  CV-D 1
9  lex.Cst  Dead 2
10 lex.Cst  CV-D 2
11 lex.Cst  Dead 3
12 lex.Cst  CV-D 3
13 lex.Xst      DM        DM
14 lex.Xst 1st CVD   1st CVD
15 lex.Xst 2nd CVD   2nd CVD
16 lex.Xst 3rd CVD    3+ CVD
17 lex.Xst  Dead 0 D(no CVD)
18 lex.Xst  CV-D 0 D(no CVD)
19 lex.Xst  Dead 1  D(1 CVD)
20 lex.Xst  CV-D 1  D(1 CVD)
21 lex.Xst  Dead 2  D(2 CVD)
22 lex.Xst  CV-D 2  D(2 CVD)
23 lex.Xst  Dead 3 D(3+ CVD)
24 lex.Xst  CV-D 3 D(3+ CVD)
```

```
> summary( L1 )
```

```
Transitions:
     To
From      DM 1st CVD 2nd CVD 3+ CVD D(no CVD) D(1 CVD) D(2 CVD) D(3+ CVD)  Records:
  DM      41      86       0      0        33        0        0          0       160
  1st CVD  0      11      48      0         0       27        0          0        86
  2nd CVD  0       0       7     24         0        0       16          0        47
  3+ CVD   0       0       0      8         0        0        0         17        25
  Sum     41      97      55     32        33       27       16         17       318

Transitions:
     To
From       Events: Risk time:  Persons:
  DM           119    1870.64       160
  1st CVD       75     342.61        86
  2nd CVD       40     112.27        47
  3+ CVD        17      92.08        25
  Sum          251    2417.60       160
```

We show the transitions separately for the two groups, using the same colouring of the states (DM, CVD, and death from each of state)

```
> clr <- c("forestgreen",heat.colors(5)[3:1])
> clr[3] <- rgb(t(col2rgb(clr[2])+col2rgb(clr[4]))/2,max=255)
> clr <- c(clr,rgb(t(col2rgb(clr)*0.6+255*0.4),max=255))
> names( clr ) <- levels(L1)
> clr
           DM       1st CVD       2nd CVD        3+ CVD     D(no CVD)      D(1 CVD)
"forestgreen"   "#FFAA00FF"     "#FF5500"   "#FF0000FF"     "#7AB97A"     "#FFCC66"
     D(2 CVD)     D(3+ CVD)
    "#FF9966"     "#FF6666"
```

```
> par( mfrow=c(1,2), mar=c(0,0,0,0) )
> for( lv in levels(L1$allocation) )
+    {
+ boxes( subset(L1,allocation==lv),
+        boxpos=list(x=c(rep(20,4),rep(80,4))+
+                    7-14*(lv=="Intensive"),
+                    y=rep(seq(90,10,,4),2) ),
+        pos.arr=c(rep(c(5,2),3),2)/10,
+        show.BE="nz", DR.sep=c(" (", ")"), hmult=1.3,
+        scale.R=100, lwd=3,
+        col.bg=clr,
+        col.border=c(clr[1:4],rep("black",4)),
+        col.txt=rep(c("white","black"),each=4) )
+ text( 10+5-10*(lv=="Intensive"), 98, lv, cex=1.5, font=2, adj=0 )
+    }
```

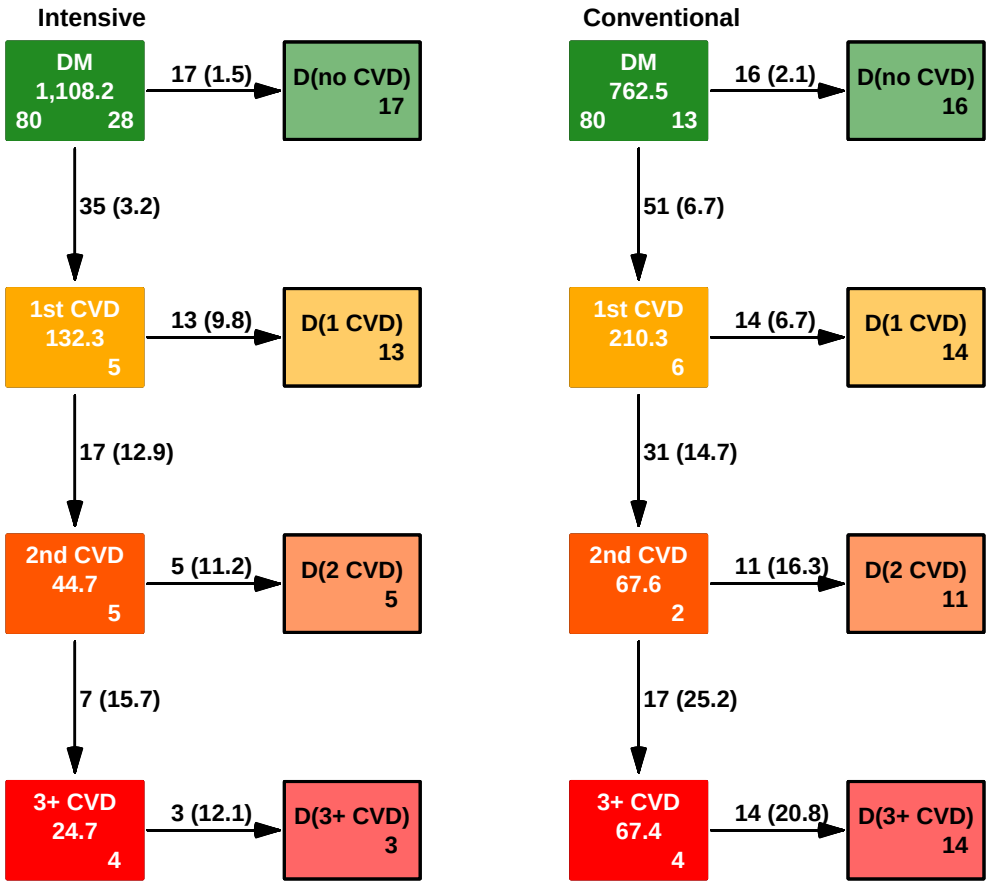

Figure 4.1: *Transitions between states of CVD and death for the two randomization groups. The number in the center of the boxes are the person-years in each state, the number at the bottom left the number of persons starting follow up in the state and the bottom right is the number of persons ending follow-up in the state. The numbers on the arrows is the number of transitions and in parenthesis the event rates per 100 PY.*

In order to answer the questions above we need a statistical models for all 7 transitions in figure 4.1.

4.2 Models for transitions

We shall model all transitions between the defined states by a function of time since baseline (baseline hazard), however only with separate effects for CVD transitions and for death. We shall moreover assume proportionality of hazards between the intervention and control groups.

The model we use in the first place has the following two separate baseline hazards:

- CVD events — this is 3 transitions which we shall assume proportional between different states of CVD progression, thus 3 proportionality parameters in addition to the baseline hazard.
- Death — this is 4 different transitions, assumed proportional between patients in different states of CVD; so three proportionality parameters on top of the baseline hazard.

Further to this model we will add a separate intervention hazard ratio for each of the 7 transitions; these will be one of the main results of the analysis.

This model will be expanded by separate sex- and linear age terms for CVD rates and mortality rates. As described the model is really two separate models; one for CVD events and one for death.

4.2.1 Modeling

We note that the transitions assumed proportional all originate in *different* states; and moreover that the mortality rates refer to *all* follow-up time.

Therefore we can do the modeling using a time-split dataset; because there are no two transitions *out* of the same box, there is no need to stack the data for analysis.

So initially we split the data in 1-month intervals of follow-up and show how the risk time and events are divided between the intervention and control groups:

```
> S1 <- splitLexis( L1, breaks=seq(0,25,1/12), time.scale="tsb" )
> summary( S1, by="allocation" )
$Intensive

Transitions:
     To
From          DM 1st CVD 2nd CVD 3+ CVD D(no CVD) D(1 CVD) D(2 CVD) D(3+ CVD)  Records:
  DM       13282      35       0      0        17        0        0         0     13334
  1st CVD      0    1592      17      0         0       13        0         0      1622
  2nd CVD      0       0     543      7         0        0        5         0       555
  3+ CVD       0       0       0    299         0        0        0         3       302
  Sum      13282    1627     560    306        17       13        5         3     15813

Transitions:
     To
From        Events: Risk time:  Persons:
  DM             52    1108.17        80
  1st CVD        30     132.29        35
  2nd CVD        12      44.70        17
  3+ CVD          3      24.70         7
  Sum            97    1309.87        80

$Conventional

Transitions:
```

```
     To
From         DM 1st CVD 2nd CVD 3+ CVD D(no CVD) D(1 CVD) D(2 CVD) D(3+ CVD)  Records:
  DM       9125      51       0      0        16        0        0          0      9192
  1st CVD     0    2528      31      0         0       14        0          0      2573
  2nd CVD     0       0     815     17         0        0       11          0       843
  3+ CVD      0       0       0    815         0        0        0         14       829
  Sum      9125    2579     846    832        16       14       11         14     13437

Transitions:
     To
From        Events: Risk time:  Persons:
  DM             67     762.47        80
  1st CVD        45     210.31        51
  2nd CVD        28      67.57        30
  3+ CVD         14      67.38        18
  Sum           154    1107.73        80
```

4.2.2 Mortality rates

We first devise a set of knots for the spline basis to use for the underlying mortality rates:

```
> levels(S1)
[1] "DM"        "1st CVD"   "2nd CVD"   "3+ CVD"    "D(no CVD)" "D(1 CVD)"  "D(2 CVD)"
[8] "D(3+ CVD)"
> ( d.kn <- c(0,with( subset(S1,lex.Xst %in% levels(lex.Xst)[5:8]),
+                     quantile(tsb+lex.dur,probs=1:4/5) ) ) )
                20%       40%       60%       80%
 0.000000  6.254346  9.174538 11.664339 16.176318
```

With these we can now fit the relevant models for mortality, the base model (`m0`) for death being the model with proportional mortalities between states of CVD and a common intervention effect.

```
> ( dlev <- levels(S1$lex.Xst)[5:8] )
[1] "D(no CVD)" "D(1 CVD)"  "D(2 CVD)"  "D(3+ CVD)"
> m0 <- glm( (lex.Xst %in% dlev ) ~
+            Ns( tsb, knots=d.kn ) + lex.Cst + allocation,
+            offset = log(lex.dur),
+            family = poisson,
+            data = S1 )
> m1  <- update( m0, . ~ . + sex + age )
> ml  <- update( m0, . ~ . - lex.Cst + I(as.integer(lex.Cst)) )
> m0i <- update( m0, . ~ . - allocation + allocation:lex.Cst )
> m1i <- update( m1, . ~ . - allocation + allocation:lex.Cst )
> mli <- update( ml, . ~ . + allocation:I(as.integer(lex.Cst)) )


> mcmp <- rbind(
+ anova( m1i, m1, m0, m0i, test="Chisq" )[-1,-(1:2)],
+ anova( mli, ml, m0, test="Chisq" )[-1,-(1:2)] )
> rownames( mcmp ) <- c("CVD*allo|s+a",
+                       "s+a|ph",
+                       "CVD*allo",
+                       "lCVD*allo",
+                       "lCVD|CVD")
> round( abs(mcmp), 3 )
             Df Deviance Pr(>Chi)
CVD*allo|s+a  3    2.711    0.438
s+a|ph        2   30.619   <2e-16
CVD*allo      3    2.392    0.495
lCVD*allo     1    0.427    0.513
lCVD|CVD      2    6.414    0.040
```

We see there is evidence that the CVD is non-linear in no of events, but also that there is no evidence that the effect of no. of CVD events is non-proportional between allocation groups.

```
> ests <- cbind( ci.exp(m1, pval=TRUE ),
+          rbind(ci.exp(ml, pval=TRUE ),NA)[c(1:5,7,8,8,6,8,8),],
+          rbind(ci.exp(m0, pval=TRUE ),NA)[c(1:9,10,10),] )
> colnames(ests)[c(1,5,9)] <- c("   m1","   ml","   m0")
> round( ests[-(1:5),], 3 )
                          m1  2.5%  97.5%     P    ml  2.5% 97.5%     P    m0  2.5%
lex.Cst1st CVD         3.076 1.821  5.194 0.000 2.082 1.728  2.51 0.000 3.721 2.206
lex.Cst2nd CVD         4.419 2.356  8.288 0.000    NA    NA    NA    NA 6.319 3.415
lex.Cst3+ CVD          7.760 4.111 14.645 0.000    NA    NA    NA    NA 8.517 4.487
allocationConventional 1.198 0.769  1.867 0.425 1.178 0.758  1.83 0.466 1.144 0.738
sexM                   1.380 0.828  2.300 0.216    NA    NA    NA    NA    NA    NA
age                    1.101 1.062  1.142 0.000    NA    NA    NA    NA    NA    NA
                        97.5%     P
lex.Cst1st CVD          6.274 0.000
lex.Cst2nd CVD         11.690 0.000
lex.Cst3+ CVD          16.168 0.000
allocationConventional  1.773 0.548
sexM                       NA    NA
age                        NA    NA
> round( cbind( ci.exp(m1 , subset=c("sex","age") ),
+               ci.exp(m1i, subset=c("sex","age") ) ), 3 )
     exp(Est.)  2.5% 97.5% exp(Est.)  2.5% 97.5%
sexM     1.380 0.828 2.300     1.333 0.797 2.230
age      1.101 1.062 1.142     1.103 1.063 1.144


> HR <- cbind( ci.exp(m1i,subset="Cst"),
+              ci.exp(m0i,subset="Cst") )
> round( rbind(HR,1/HR[4:7,,drop=FALSE]), 3 )
                                    exp(Est.)  2.5%  97.5% exp(Est.)  2.5%  97.5%
lex.Cst1st CVD                          4.813 2.323  9.971     5.564 2.697 11.482
lex.Cst2nd CVD                          4.678 1.693 12.924     6.334 2.288 17.536
lex.Cst3+ CVD                           8.023 2.298 28.005     7.775 2.199 27.493
lex.CstDM:allocationConventional        1.664 0.835  3.315     1.487 0.749  2.951
lex.Cst1st CVD:allocationConventional   0.712 0.332  1.529     0.699 0.328  1.487
lex.Cst2nd CVD:allocationConventional   1.371 0.457  4.115     1.356 0.452  4.070
lex.Cst3+ CVD:allocationConventional    1.348 0.377  4.819     1.457 0.408  5.201
lex.CstDM:allocationConventional        0.601 1.197  0.302     0.673 1.335  0.339
lex.Cst1st CVD:allocationConventional   1.404 3.012  0.654     1.431 3.045  0.673
lex.Cst2nd CVD:allocationConventional   0.729 2.189  0.243     0.737 2.212  0.246
lex.Cst3+ CVD:allocationConventional    0.742 2.652  0.208     0.686 2.450  0.192
> HR <- cbind( ci.exp(m1 ,subset=c("Cst","allo"), pval=TRUE ),
+              ci.exp(m0 ,subset=c("Cst","allo"), pval=TRUE ) )
> IHR <- 1/HR[4,,drop=F]
> IHR[,c(4,8)] <- 1/IHR[,c(4,8)]
> round( rbind(HR,IHR), 3 )
                       exp(Est.)  2.5%  97.5%     P exp(Est.)  2.5%  97.5%     P
lex.Cst1st CVD             3.076 1.821  5.194 0.000     3.721 2.206  6.274 0.000
lex.Cst2nd CVD             4.419 2.356  8.288 0.000     6.319 3.415 11.690 0.000
lex.Cst3+ CVD              7.760 4.111 14.645 0.000     8.517 4.487 16.168 0.000
allocationConventional     1.198 0.769  1.867 0.425     1.144 0.738  1.773 0.548
allocationConventional     0.835 1.301  0.536 0.425     0.874 1.355  0.564 0.548
```

We then update the HR collector:

```
> load( file="../data/mainCI.Rda" )
> dimnames( mainCI )[[1]][5]
[1] "Death | CVD state"
```

```
> mainCI[5,,] <- rbind(IHR[c(1,3,2,4)],IHR[c(1,3,2,4)+4])
> save( mainCI, file="../data/mainCI.Rda" )
```

It is seen that mortality increases by CVD status, as well as by age. Also it is seen that the CVD-effects are smaller in the model where age is accounted for; this is because the persons with more advanced CVD status are older. The HR between intervention and control is 0.83 (0.53;1.30) when controlled for sex and age.

A graphical overview of the estimates of the intervention effect:

```
> e0  <- ci.exp(m0 ,subset="allocation")
> e1  <- ci.exp(m1 ,subset="allocation")
> e0i <- ci.exp(m0i,subset="allocation")
> e1i <- ci.exp(m1i,subset="allocation")
> rownames( e0  ) <- "Overall"
> rownames( e0i ) <- gsub( "lex.Cst", "",
+                          gsub( ":allocationConventional", "", rownames(e0i) ) )
> rownames( e1i ) <- gsub( "lex.Cst", "",
+                          gsub( ":allocationConventional", "", rownames(e1i) ) )
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6 )
> yp <- c(5,3:0)
>  plotEst( 1/rbind(e0,e0i), txtpos=yp, y=yp+0.1,
+           vref=1, xlog=TRUE, xlim=c(1/5,5), grid=c(2:10/10,1.5,2:5),
+           xlab="Mortality HR: Intensive vs. conventional", cex=2, lwd=5 )
> linesEst( 1/rbind(e1,e1i), y=yp-0.1, col="gray", cex=2, lwd=5 )
```

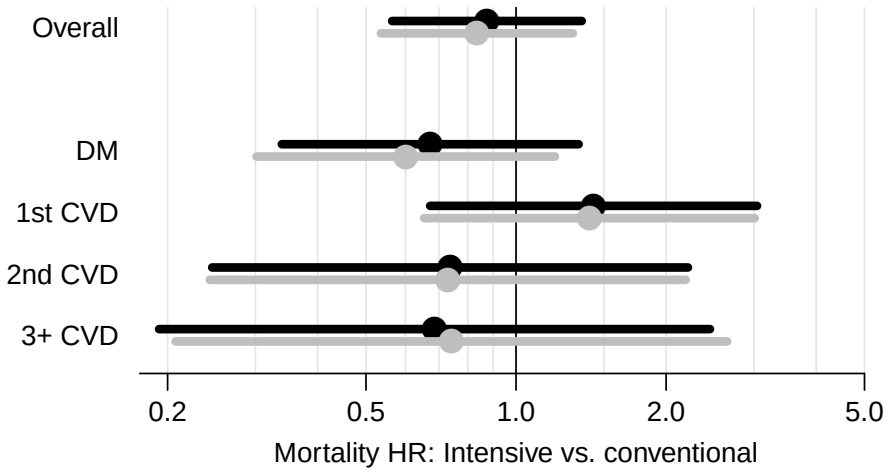

Figure 4.2: *Ratio of mortality between intensive and conventional group, in different states of CVD progression. Gray estimates are from models with sex and age effect.*

From figure 4.2 it is seen that there is no significant difference in mortality between the two groups when we condition on the CVD status. There seems to be a slightly `higher` mortality in the intervention group among persons with *one* CVD event than among persons with none or 2+ CVD events. As it is neither statistically significant nor biologically plausible, we shall use the proportional hazards approach; that is the one where we assume that intervention has the same effect (as measured by the HR) on all transition rates — the top estimate in the graph.

4.2.3 CVD rates

When we model the CVD occurrence rates (hazards) we must exclude the follow-up time in the state "`3+ CVD`" because these are not at risk of any further CVD event. Moreover the

specification of the events are a bit more tricky, because these are not just the exit status being equal to one of the CVD states, but *also* that `lex.Xst` is different from `lex.Cst`, because some of the split records will have both `lex.Cst` and `lex.Xst` equal to a given CVD state. Otherwise the analyses are along the same lines as the mortality analyses:

```
> ( clev <- levels(S1$lex.Xst)[2:4] )
[1] "1st CVD" "2nd CVD" "3+ CVD"

> c0 <- glm( ( (lex.Xst %in% clev) & (lex.Cst!=lex.Xst) ) ~
+            Ns( tsb, knots=d.kn ) + lex.Cst + allocation,
+            offset = log(lex.dur),
+            family = poisson,
+            data = subset( S1, lex.Cst!="3+ CVD" ) )
> c1  <- update( c0, . ~ . + sex + age )
> cl  <- update( c0, . ~ . - lex.Cst + I(as.integer(lex.Cst)) )
> c0i <- update( c0, . ~ . - allocation + allocation:lex.Cst )
> c1i <- update( c1, . ~ . - allocation + allocation:lex.Cst )
> cli <- update( cl, . ~ . + allocation:I(as.integer(lex.Cst)) )
> ccmp <- rbind(
+ anova( c1i, c1, c0, c0i, test="Chisq" )[-1,-(1:2)],
+ anova( cli, cl, c0, test="Chisq" )[-1,-(1:2)] )
> rownames( ccmp ) <- c("CVD*allo|s+a",
+                       "s+a|ph",
+                       "CVD*allo",
+                       "lCVD*allo",
+                       "lCVD|CVD")
> round( abs(ccmp), 3 )
             Df Deviance Pr(>Chi)
CVD*allo|s+a  2    2.685    0.261
s+a|ph        2   12.538    0.002
CVD*allo      2    2.712    0.258
lCVD*allo     1    0.988    0.320
lCVD|CVD      1    2.259    0.133

> cest <- cbind( ci.exp(c1, pval=TRUE ),
+          rbind(ci.exp(cl, pval=TRUE ),NA)[c(1:5,7,8,6,8,8),],
+          rbind(ci.exp(c0, pval=TRUE ),NA)[c(1:8,9,9),] )
> colnames(cest)[c(1,5,9)] <- c("   c1","   cl","   c0")
> round( cest[-(1:5),], 3 )
                           c1  2.5% 97.5%     P    cl  2.5% 97.5%     P    c0  2.5% 97.5%
lex.Cst1st CVD          2.427 1.671 3.523 0.000 2.167 1.749 2.684 0.000 2.740 1.894 3.964
lex.Cst2nd CVD          3.481 2.148 5.643 0.000    NA    NA    NA    NA 4.260 2.661 6.818
allocationConventional 1.813 1.296 2.536 0.001 1.788 1.282 2.493 0.001 1.738 1.245 2.426
sexM                    1.325 0.897 1.958 0.157    NA    NA    NA    NA    NA    NA    NA
age                     1.042 1.016 1.068 0.001    NA    NA    NA    NA    NA    NA    NA
                             P
lex.Cst1st CVD          0.000
lex.Cst2nd CVD          0.000
allocationConventional 0.001
sexM                       NA
age                        NA

> round( cbind( ci.exp(c1 , subset=c("sex","age") ),
+               ci.exp(c1i, subset=c("sex","age") ) ), 3 )
     exp(Est.)  2.5% 97.5% exp(Est.)  2.5% 97.5%
sexM     1.325 0.897 1.958     1.296 0.876 1.919
age      1.042 1.016 1.068     1.043 1.017 1.069
```

From the models we can also extract both the HRs and the inverse HRs corresponding to intervention versus control — the latter are in the 3 last lines here:

```
> HR <- cbind( ci.exp(c1i,subset="Cst"),
+              ci.exp(c0i,subset="Cst") )
> round( rbind(HR,1/HR[3:5,,drop=FALSE]), 3 )
```

```
                                    exp(Est.)  2.5% 97.5% exp(Est.)  2.5%  97.5%
lex.Cst1st CVD                          3.625 2.016 6.519     4.101 2.288  7.350
lex.Cst2nd CVD                          3.929 1.703 9.068     4.722 2.053 10.863
lex.CstDM:allocationConventional        2.229 1.443 3.444     2.134 1.383  3.291
lex.Cst1st CVD:allocationConventional   1.195 0.658 2.170     1.143 0.632  2.066
lex.Cst2nd CVD:allocationConventional   1.799 0.719 4.497     1.763 0.705  4.411
lex.CstDM:allocationConventional        0.449 0.693 0.290     0.469 0.723  0.304
lex.Cst1st CVD:allocationConventional   0.837 1.519 0.461     0.875 1.581  0.484
lex.Cst2nd CVD:allocationConventional   0.556 1.390 0.222     0.567 1.419  0.227
```

```
> HR <- cbind( ci.exp(c1, subset=c("Cst","allo"), pval=TRUE),
+              ci.exp(c0, subset=c("Cst","allo"), pval=TRUE) )
> round( rbind(HR,1/HR[3,,drop=FALSE]), 3 )
```

```
                      exp(Est.)  2.5% 97.5%        P exp(Est.)  2.5% 97.5%       P
lex.Cst1st CVD            2.427 1.671 3.523    0.000     2.740 1.894 3.964   0.000
lex.Cst2nd CVD            3.481 2.148 5.643    0.000     4.260 2.661 6.818   0.000
allocationConventional    1.813 1.296 2.536    0.001     1.738 1.245 2.426   0.001
allocationConventional    0.552 0.772 0.394 1934.702     0.575 0.803 0.412 863.382
```

```
> IHR <- 1/HR[3,,drop=F]
> IHR[,c(4,8)] <- 1/IHR[,c(4,8)]
> round( rbind(HR,IHR), 3 )
```

```
                      exp(Est.)  2.5% 97.5%     P exp(Est.)  2.5% 97.5%     P
lex.Cst1st CVD            2.427 1.671 3.523 0.000     2.740 1.894 3.964 0.000
lex.Cst2nd CVD            3.481 2.148 5.643 0.000     4.260 2.661 6.818 0.000
allocationConventional    1.813 1.296 2.536 0.001     1.738 1.245 2.426 0.001
allocationConventional    0.552 0.772 0.394 0.001     0.575 0.803 0.412 0.001
```

We then update the HR collector:

```
> load( file="../data/mainCI.Rda" )
> dimnames( mainCI )[[1]][6]
[1] "CVD event | CVD state"

> mainCI[6,,] <- rbind(IHR[c(1,3,2,4)],IHR[c(1,3,2,4)+4])
> save( mainCI, file="../data/mainCI.Rda" )
```

As before we see smaller effect of the intervention after the first CVD; formally the intervention effect on CVD occurrence is non-significant among persons with one or two post-baseline CVD, but there is no consistently declining effect as shown in figure 4.3:

```
> e0  <- ci.exp(c0 ,subset="allocation")
> e1  <- ci.exp(c1 ,subset="allocation")
> e0i <- ci.exp(c0i,subset="allocation")
> e1i <- ci.exp(c1i,subset="allocation")
> rownames( e0  ) <- "Overall"
> rownames( e0i ) <- gsub( "lex.Cst", "",
+                          gsub( ":allocationConventional", "", rownames(e0i) ) )
> rownames( e1i ) <- gsub( "lex.Cst", "",
+                          gsub( ":allocationConventional", "", rownames(e1i) ) )
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6 )
> yp <- c(4,2:0)
>  plotEst( 1/rbind(e0,e0i), txtpos=yp, y=yp+0.1,
+           vref=1, xlog=TRUE, xlim=c(1/5,5), grid=c(2:10/10,1.5,2:5),
+           xlab="Invervention vs. conventional CVD HR", cex=2, lwd=5 )
> linesEst( 1/rbind(e1,e1i), y=yp-0.1, col="gray", cex=2, lwd=5 )
```

As for mortality rates, for CVD rates there is no convincing evidence of interaction between allocation and state, so we use the model with common intervention HR for all CVD event rates.

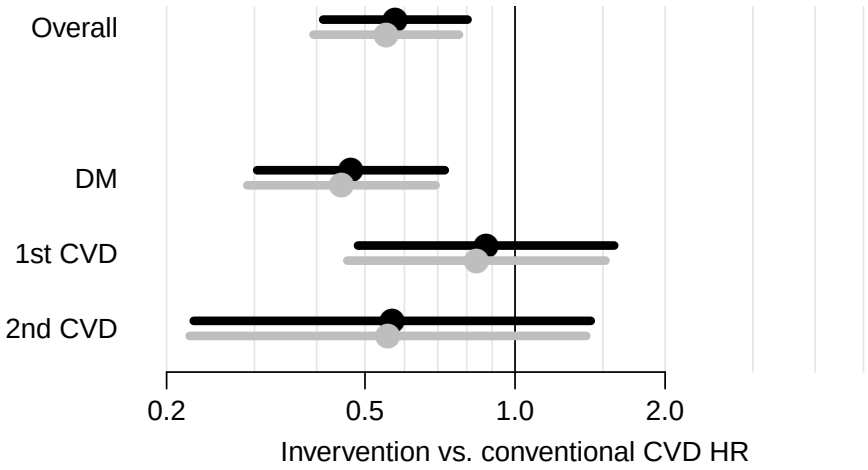

Figure 4.3: *Ratio of mortality between intervention and conventional group, in different states of CVD progression. The gray estimates are from a model with conventional for age and sex.*

4.2.4 Summary of intervention

In summary, we have thus two intervention effects; one for CVD-rates and the other for mortality rates — both of which are independent of the current CVD status of the patients (0, 1, 2 (or 3) CVD events).

For completeness we have also included HR from the analysis of overall mortality where CVD status was ignored:

Transition	Hazard ratio (95% c.i)				
	Conv vs. Int		Int vs. Conv		P
Crude					
Mortality	1.82	(1.20;2.75)	0.55	(0.36;0.83)	0.005
Mort|CVD	1.14	(0.74;1.77)	0.87	(0.56;1.36)	0.548
CVD	1.74	(1.25;2.43)	0.58	(0.41;0.80)	0.001
Controlled for age and sex					
Mortality	1.90	(1.25;2.89)	0.53	(0.35;0.80)	0.003
Mort|CVD	1.20	(0.77;1.87)	0.83	(0.54;1.30)	0.425
CVD	1.81	(1.30;2.54)	0.55	(0.39;0.77)	0.001

Note that the effect of intervention on mortality given CVD status (Mort|CVD) is formally not significant; this is because the effect shown is the effect on the *conditional* mortality *given* CVD state — the 4 horizontal transitions shown in figure 4.1. The intervention is only significant for the overall mortality *disregarding* CVD status.

Thus it seems that the major effect of the intervention on mortality is mediated by (prevention of) CVD occurrence.

4.3 Predictions

We now have (several) statistical models for all 7 transitions shown in figure 4.1 and this enables us to estimate the state occupancy probabilities for a given base population under these models. With state occupancy probabilities estimated as continuous functions of time we can compute the fraction of persons that see 0, 1, 2 or 3 CVDs during the first 20 years of follow-up; and we can compute the years spent alive with 0, 1, 2 and 3 CVD events. These quantities can then be compared between the persons in the two groups.

Both the mortality models and the CVD models we have fitted are cross-classified by whether they include a sex and an age term and whether they include an interaction between CVD state and intervention.

There are two substantially different scenarios we want to address:

- How does the follow-up look for the Steno2 population, and how is it influenced by the intervention?

  This can be addressed using either the models with or without age and sex-effects, since the prediction anyway will be for a population with the same age- and sex-composition as the Steno2 population.

- How does the follow-up look for T2D patients of a given sex and age?

  This can only be addressed using the models with age and sex-effects, since the prediction otherwise will be the same regardless of the population's age- and sex-composition.

In view of this it seems most sensible to use the models with age- and sex effects as basis for all predictions. Also in the light of the very significant effect of these two variables on the rates. Note that even if we have a randomized study, the fact that CVD rates are influenced by age and sex implies that the age- and sex-distribution in the CVD-states is not comparable between allocation groups.

There is no formal evidence in data that that neither the mortality nor the CVD HRs between intervention and conventional vary across CVD states:

```
> round( cbind( rbind( ci.exp(m1 , subset="allo", pval=T),
+                      ci.exp(m1i, subset="allo", pval=T) ),
+               rbind( ci.exp(c1 , subset="allo", pval=T),
+                      ci.exp(c1i, subset="allo", pval=T), NA ) ), 3 )
                                   exp(Est.)  2.5% 97.5%     P exp(Est.)  2.5% 97.5%
allocationConventional                 1.198 0.769 1.867 0.425     1.813 1.296 2.536
lex.CstDM:allocationConventional       1.664 0.835 3.315 0.148     2.229 1.443 3.444
lex.Cst1st CVD:allocationConventional  0.712 0.332 1.529 0.384     1.195 0.658 2.170
lex.Cst2nd CVD:allocationConventional  1.371 0.457 4.115 0.574     1.799 0.719 4.497
lex.Cst3+ CVD:allocationConventional   1.348 0.377 4.819 0.646        NA    NA    NA
                                       P
allocationConventional                 0.001
lex.CstDM:allocationConventional       0.000
lex.Cst1st CVD:allocationConventional  0.558
lex.Cst2nd CVD:allocationConventional  0.209
lex.Cst3+ CVD:allocationConventional      NA
> anova( m1, m1i, test="Chisq")
Analysis of Deviance Table

Model 1: (lex.Xst %in% dlev) ~ Ns(tsb, knots = d.kn) + lex.Cst + allocation +
    sex + age
Model 2: (lex.Xst %in% dlev) ~ Ns(tsb, knots = d.kn) + lex.Cst + sex +
```

```
    age + lex.Cst:allocation
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1     29239     1145.1
2     29236     1142.4  3   2.7111   0.4383
> anova( c1, c1i, test="Chisq")
Analysis of Deviance Table

Model 1: ((lex.Xst %in% clev) & (lex.Cst != lex.Xst)) ~ Ns(tsb, knots = d.kn) +
    lex.Cst + allocation + sex + age
Model 2: ((lex.Xst %in% clev) & (lex.Cst != lex.Xst)) ~ Ns(tsb, knots = d.kn) +
    lex.Cst + sex + age + lex.Cst:allocation
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1     28109     1833.3
2     28107     1830.6  2   2.6854   0.2611
```

Also, the deviance statistics are both smaller then 3.84, so it can be safely excluded that the non-significance is a mere product of inflation of degrees of freedom. Hence we shall make predictions based on the models with inervention effects constant across CVD strata.

4.3.1 PH model predictions

In order to set up a prediction scenario we need a transition object, that is a list of lists indicating what models to be used for each of the transitions. We set up two transition objects (models for transition intensities), one using the same proportionality between randomization groups for all transitions of each type (death / CVD) and th eother allowing differential allocation effects on rates across CVD states. Boteh sets of models used are however proportional hazards models along the time since baseline scale.

```
> levels( S1 )
[1] "DM"        "1st CVD"   "2nd CVD"   "3+ CVD"    "D(no CVD)" "D(1 CVD)"  "D(2 CVD)"
[8] "D(3+ CVD)"
> TM1 <- list( "DM" = list( "D(no CVD)" = m1,
+                           "1st CVD" = c1 ),
+         "1st CVD" = list(  "D(1 CVD)" = m1,
+                           "2nd CVD" = c1 ),
+         "2nd CVD" = list(  "D(2 CVD)" = m1,
+                            "3+ CVD" = c1 ),
+          "3+ CVD" = list( "D(3+ CVD)" = m1 ) )
> TM1i<- list( "DM" = list( "D(no CVD)" = m1i,
+                           "1st CVD" = c1i ),
+         "1st CVD" = list(  "D(1 CVD)" = m1i,
+                           "2nd CVD" = c1i ),
+         "2nd CVD" = list(  "D(2 CVD)" = m1i,
+                            "3+ CVD" = c1i ),
+          "3+ CVD" = list( "D(3+ CVD)" = m1i ) )
> lapply( TM1, names )
$DM
[1] "D(no CVD)" "1st CVD"

$`1st CVD`
[1] "D(1 CVD)" "2nd CVD"

$`2nd CVD`
[1] "D(2 CVD)" "3+ CVD"

$`3+ CVD`
[1] "D(3+ CVD)"
```

Note that there is only two models used in each of the transition objects; because we are using the proportional hazards assumption for the CVD rates and for the mortality rates separately.

We need an initial dataset from which to simulate; that is an assumption about the distribution of covariates among persons that we use for simulation of follow-up.

4.3.2 The Steno2 population

In the case of the Steno2 population, we devise an initial dataset with *all* persons in the study in two copies; one assigned to intervention and one to conventional. Note that we use "`subset`" in order to keep attributes of the Lexis object `L0`, and that that we explicitly indicate the rows when we assign new values in order to keep the attributes (factor levels, that is) of the variables from `L0`:

```
> St2 <- subset( L0, select = c(timeScales(L1),"lex.Cst","lex.dur",
+                               "allocation","sex") )
> levels(St2$lex.Cst) <- levels(S1$lex.Cst)
> nr <- nrow(St2)
> St2[1:nr,"tsb"] <- 0
> St2[1:nr,"lex.Cst"] <- levels(St2$lex.Cst)[1]
> St2[1:nr,"allocation"] <- levels(S1$allocation)[1]
> St2 <- rbind( St2, St2 )
> St2[1:nr,"allocation"] <- levels(S1$allocation)[2]
> str( St2 )
Classes 'Lexis' and 'data.frame':        320 obs. of  8 variables:
 $ per       : num  1993 1993 1993 1993 1993 ...
 $ age       : num  61.1 46.5 50 48.6 57.2 ...
 $ dur       : num  2.26 11.23 10.8 16.57 6.82 ...
 $ tsb       : num  0 0 0 0 0 0 0 0 0 0 ...
 $ lex.Cst   : Factor w/ 8 levels "DM","1st CVD",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ lex.dur   : num  21.42 21.23 8.28 10.08 4.88 ...
 $ allocation: Factor w/ 2 levels "Intensive","Conventional": 2 2 2 2 2 2 2 2 2 2 ...
 $ sex       : Factor w/ 2 levels "F","M": 2 2 2 2 2 2 1 2 2 2 ...
 - attr(*, "breaks")=List of 4
  ..$ per: NULL
  ..$ age: NULL
  ..$ dur: NULL
  ..$ tsb: NULL
 - attr(*, "time.scales")= chr  "per" "age" "dur" "tsb"
 - attr(*, "time.since")= chr  "" "" "" ""
> with( St2, table(sex,allocation) )
   allocation
sex Intensive Conventional
  F        41           41
  M       119          119
```

The data frame we now have for prediction has 2 copies of each person from the Steno2 study, one allocated to "Intensive" and one allocated to "Conventional". When we simulate we will then get predicted results for the entire Steno2 population with the sex- and age-distribution af *all* patients in the study *as if* they were all allocated to each of the treatments.

```
> system.time( sim2 <- simLexis( Tr = TM1,
+                                init = St2,
+                                   N = 100,
+                           time.pts = seq(0,21,0.1) ) )
   user  system elapsed
125.642   7.101 132.728
> system.time( sim2i<- simLexis( Tr = TM1i,
+                                init = St2,
+                                   N = 100,
+                           time.pts = seq(0,21,0.1) ) )
```

```
   user  system elapsed
124.118   3.743 129.355
> save( sim2, sim2i, file="../data/sim2.Rda" )
```

The Lexis objects `sim2` and `sim2i` now contain a simulated follow up of a cohort of 32,000 persons; 100 copies of each of the original Steno2 participants (that is of the age and sex of the persons) assigned to conventional and 100 copies assigned to intervention. Thus the observed state distribution in each of the two subsets will therefore be an estimate of the state probability distribution induced by the estimated transition intensities: Further, after simulation of the temporal distribution of state occupancy, we will show the distribution across states in a certain *order* of the states:

```
> prm <- c(1:4,8:5)
> levels( St2$lex.Cst )[prm]
[1] "DM"        "1st CVD"   "2nd CVD"   "3+ CVD"    "D(3+ CVD)" "D(2 CVD)"  "D(1 CVD)"
[8] "D(no CVD)"


> load( file="../data/sim2.Rda" )
> summary( sim2, by="allocation" )
$Intensive

Transitions:
     To
From         DM 1st CVD 2nd CVD 3+ CVD D(no CVD) D(1 CVD) D(2 CVD) D(3+ CVD)  Records:
  DM       4970    7374       0      0      3656        0        0          0     16000
  1st CVD     0    1593    3263      0         0     2518        0          0      7374
  2nd CVD     0       0     678   1425         0        0     1160          0      3263
  3+ CVD      0       0       0    477         0        0        0        948      1425
  Sum      4970    8967    3941   1902      3656     2518     1160        948     28062

Transitions:
     To
From        Events: Risk time:  Persons:
  DM          11030  206419.26     16000
  1st CVD      5781   36158.20      7374
  2nd CVD      2585   10399.24      3263
  3+ CVD        948    4648.26      1425
  Sum         20344  257624.96     16000

$Conventional

Transitions:
     To
From         DM 1st CVD 2nd CVD 3+ CVD D(no CVD) D(1 CVD) D(2 CVD) D(3+ CVD)  Records:
  DM       2647   10261       0      0      3092        0        0          0     16000
  1st CVD     0    1115    6297      0         0     2849        0          0     10261
  2nd CVD     0       0     675   3747         0        0     1875          0      6297
  3+ CVD      0       0       0   1030         0        0        0       2717      3747
  Sum      2647   11376    6972   4777      3092     2849     1875       2717     36305

Transitions:
     To
From        Events: Risk time:  Persons:
  DM          13353  162159.43     16000
  1st CVD      9146   39737.00     10261
  2nd CVD      5622   15918.40      6297
  3+ CVD       2717   12806.19      3747
  Sum         30838  230621.02     16000
> system.time(
+ StC <- nState( subset(sim2,allocation=="Conventional"),
+                at=seq(0,20,0.1), from=0, time.scale="tsb" ) )
```

```
   user  system elapsed
 13.727   0.000  13.724
> StI <- nState( subset(sim2,allocation=="Intensive"),
+                at=seq(0,20,0.1), from=0, time.scale="tsb" )
> pC <- pState( StC, perm=prm )
> pI <- pState( StI, perm=prm )
> system.time(
+ StCi <- nState( subset(sim2i,allocation=="Conventional"),
+                 at=seq(0,20,0.1), from=0, time.scale="tsb" ) )
   user  system elapsed
 11.942   0.000  11.936
> StIi <- nState( subset(sim2i,allocation=="Intensive"),
+                 at=seq(0,20,0.1), from=0, time.scale="tsb" )
> pCi <- pState( StCi, perm=prm )
> pIi <- pState( StIi, perm=prm )
> save( pC, pI, pCi, pIi, file="../data/pSt2.Rda" )

> load( file="../data/pSt2.Rda" )
> head( pC )
     State
when         DM   1st CVD   2nd CVD    3+ CVD D(3+ CVD)  D(2 CVD)  D(1 CVD) D(no CVD)
  0   1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000         1
  0.1 0.9925000 0.9993125 0.9993125 0.9993750 0.9993750 0.9993750 0.9993750         1
  0.2 0.9848125 0.9988125 0.9988750 0.9989375 0.9989375 0.9989375 0.9990000         1
  0.3 0.9795000 0.9980625 0.9985000 0.9986250 0.9986250 0.9986250 0.9988125         1
  0.4 0.9727500 0.9971875 0.9981250 0.9982500 0.9982500 0.9982500 0.9984375         1
  0.5 0.9652500 0.9963125 0.9976875 0.9979375 0.9979375 0.9979375 0.9981875         1

> ppp <- function( pI=pI, pC=pC, rv=FALSE, ln=FALSE)
+    {
+ par( mfrow=c(1,2), mar=c(0,3,0,0), oma=c(3,0,2,3-2*rv), las=1 )
+
+ (c50I <- approx(pI[,1],as.numeric(rownames(pI)),0.5)$y)
+ (c50C <- approx(pC[,1],as.numeric(rownames(pI)),0.5)$y)
+ (d50C <- approx(pC[,4],as.numeric(rownames(pI)),0.5)$y)
+
+ plot( pI, col=clr[prm], xlab="" )
+ lines( as.numeric(rownames(pI)), pI[,4], col="black", lwd=4 )
+ if( ln ) {
+      abline( h=0.5, col="gray" )
+      segments(c50I,0,c50I,0.5,col="gray")
+          }
+ axis( side=4, at=0:5/5 )
+ axis( side=4, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4, at=0:100  /100, tcl=-0.2, labels=NA )
+ mtext( "Intensive", side=3, line=1 )
+
+ plot( pC, col=clr[prm], xlim=if(rv) c(20,0) else c(0,20),
+       xlab="", yaxt="n", ylab="", col.lab="transparent" )
+ lines( as.numeric(rownames(pC)), pC[,4], col="black", lwd=4 )
+ if( ln ) {
+    abline(h=0.50,col="gray")
+    segments(c50C,0,c50C,0.5,col="gray")
+    segments(d50C,0,d50C,0.5,col="gray")
+          }
+ axis( side=4-2*rv, at=1:5/5, if(rv) labels=NA )
+ axis( side=4-2*rv, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4-2*rv, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4-2*rv, at=0:100  /100, tcl=-0.2, labels=NA )
+ mtext( "Conventional", side=3, line=1 )
+
+ mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
+ }
> ppp(pI,pC,ln=TRUE,rv=TRUE)
```

```
> ppp(pIi,pCi,rv=TRUE,ln=TRUE)


> postscript( file="../graph/Fig1.eps", height=6, width=8 )
> ppp(pI,pC)
> dev.off()
null device
          1
> postscript( file="../graph/Fig1r.eps", height=6, width=8 )
> ppp(pI,pC,rv=TRUE)
> dev.off()
null device
          1


> ppp(pIi,pCi,rv=TRUE)
```

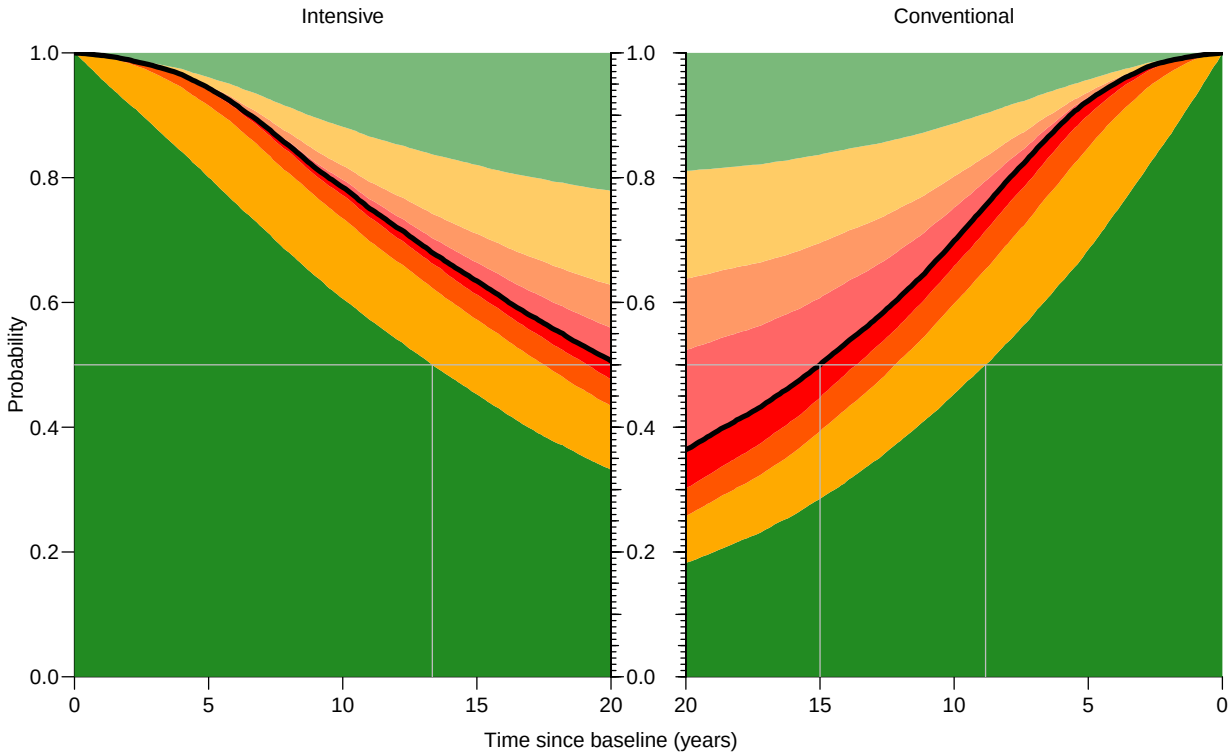

Figure 4.4: *State distribution for a patient population with the same age- and sex-distribution as the Steno 2 population using models with constant intervention effect across CVD states. The colors of the states correspond to the colouring in figure 4.1.*

4.4 Age- and sex specific results

We also want to make predictions conditional on the age at entry; and separately for men and omen, because these two variables have major impact on both mortality and CVD-rates, and in particular because these are more likely to be generalizable — the overall figures derived above are essentially only valid for a patient population of the exact same age and sex composition as the Steno 2 population.

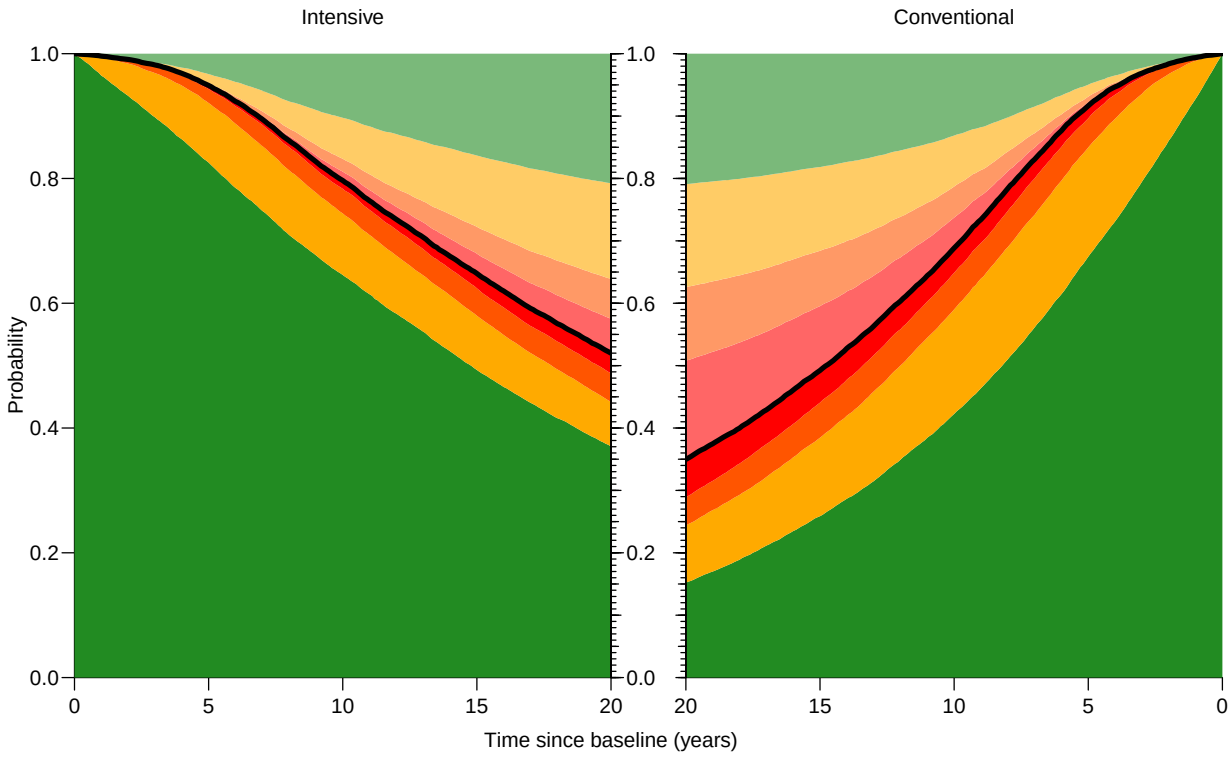

Figure 4.5: *State distribution for a patient population with the same age- and sex-distribution as the Steno 2 population using models with constant intervention effect across CVD states. The colors of the states correspond to the colouring in figure 4.1. This is the same plot as in figure 4.4, except that the time axis is reversed for the conventional group.*

Thus we simulate the life-course for a sample of 10,000 identical patients for each combination of allocation, sex and 5 different ages at entry (45, 50, 55, 60, 65) covering the range of entry age to the Steno2 study.

The results in the form of `pState` objects will be stored in an array, that is a list with a `dim` attribute of length 3, corresponding to the three variables (allocation, age and sex). The elements of the list is then filled in a loop over the dimensions:

```
> lp <- list()
> length( lp ) <- 20
> dim( lp ) <- c(5,2,2)
> dimnames( lp ) <- list( age = seq(45,65,5),
+                         sex = levels(L0$sex),
+                         grp = levels(L0$allocation) )
```

We make a copy of this to hold the results from simulations from the models with varying intervention effects:

```
> lpi <- lp
```

Then we can fill the two lists with `pState` objects from simulations, but we first need to define a starting Lexis object (a person with characteristics that we want to simulate). All persons are taken to start at 0 (since baseline) and in the DM state. Since the Lexis object also contains the timescales per and dur, we must define values fr these too, but the particular values chosen for these will have no effect om the results, because the models for the transitions do not depend on these.

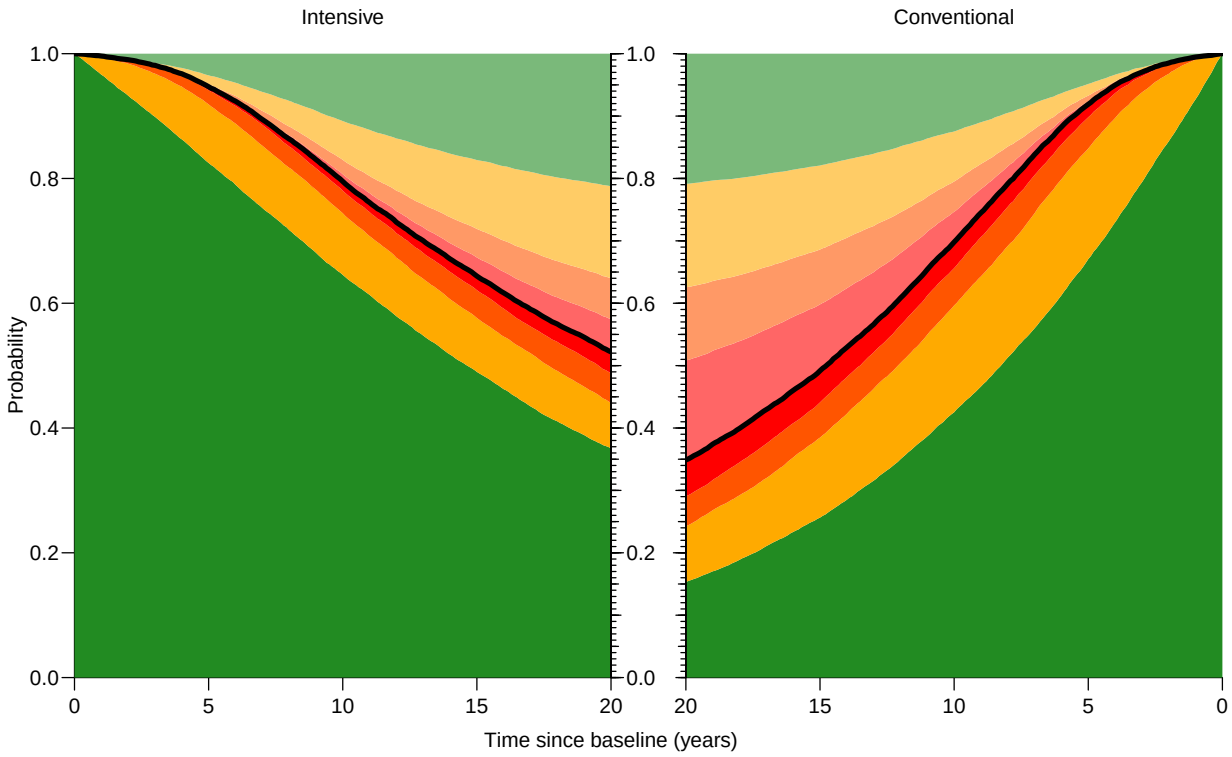

Figure 4.6: *State distribution for a patient population with the same age- and sex-distribution as the Steno 2 population using models with different intervention effect across CVD states. The colors of the states correspond to the colouring in figure 4.1.*

```
> Sb <- St2[NULL,]
> Sb[1,"tsb"] <- 0
> Sb[1,"per"] <- 1993
> Sb[1,"dur"] <- 5
> Sb[1,"lex.Cst"] <- "DM"
```

Once this is in place, we can simulate and compute the state distribution at different levels; note that we specify the starting age, sex and allocation inside the loop, since these are the values of the loop variables:

```
> system.time(
+ for( ig in dimnames(lp)[[3]] )
+ for( ia in dimnames(lp)[[1]] )
+ for( is in dimnames(lp)[[2]] )
+    {
+ Sb[1,"age"] <- as.numeric(ia)
+ Sb[1,"sex"] <- is
+ Sb[1,"allocation"] <- ig
+ stp <- simLexis( Tr = TM1,
+                  init = Sb,
+                     N = 5000,
+             time.pts = seq(0,21,0.1) )
+ nst <- nState( stp, at=seq(0,20,0.1), from=0, time.scale="tsb" )
+ pst <- pState( nst, perm=prm )
+ lp[[ia,is,ig]] <- pst
+ stp <- simLexis( Tr = TM1i,
+                  init = Sb,
+                     N = 5000,
+             time.pts = seq(0,21,0.1) )
+ nst <- nState( stp, at=seq(0,20,0.1), from=0, time.scale="tsb" )
+ pst <- pState( nst, perm=prm )
```

```
+ lpi[[ia,is,ig]] <- pst
+    } )
   user  system elapsed
874.591   0.029 874.463
> save( lp, lpi, file="../data/lp.Rda" )

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

The object `lp` is now a list with 3 dimensions (intervention, age, sex), where each element is a `pState` object, so that we can refer easily to results for any combination of age, sex and allocation.

4.4.1 State distribution

Using the `lp` and `lpi` objects we can now compare the state distribution between persons allocated to intervention and conventional, separately for all ten combinations of age (45, 50, 55, 60, 65) and sex (M, F).

Basically it is the same graph as for the Steno2 population

```
> pl10 <- function(lp){
+ par( mfrow=c(5,4), mar=c(0,0,0,3), oma=c(4,4,4,0), las=1 )
+ for( ia in dimnames(lp)[["age"]] )
+ for( is in dimnames(lp)[["sex"]][2:1] )
+ for( ig in dimnames(lp)[["grp"]] )
+    {
+ oax <- ( ig==dimnames(lp)[["grp"]][1] )
+ ptmp <- lp[[ia,is,ig]]
+ plot( ptmp, col=clr[prm], xaxt="n", yaxt="n",
+       xlim=if( oax ) c(0,20) else c(20,0) )
+ if( ia=="45") mtext( ig, side=3, line=1, at=10, outer=FALSE )
+ if( is=="M" & ig=="Intensive" )
+   {
+   axis( side=4-2*oax, at=1:5/5 )
+ # mtext( ia, side=2, line=3, outer=FALSE )
+   }
+ if( is=="F" & ig=="Conventional" ) mtext(ia,side=4,at=0.5,line=0.5,las=1)
+ axis( side=2, labels=NA )
+ if( ia=="65")
+   {
+   axis( side=1 )
+   axis( side=4, at=0, labels="0.0" )
+   if( ia=="65" & is=="M" & oax ) axis( side=2, at=0, labels="0.0" )
+   }
+ lines( as.numeric(rownames(ptmp)), ptmp[,4], col="black", lwd=4 )
+ # text( 1, 0.05, paste( is, ", ", ia, "\n", ig, esp="" ),
+ #       adj=c(1-oax,0), font=2, col="white", cex=1.5 )
+ axis( side=2+2*oax, at=1:5*20 /100,  if(oax)  labels=NA )
+ axis( side=2+2*oax, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=2+2*oax, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=2+2*oax, at=0:100  /100, tcl=-0.2, labels=NA )
+    }
+ #mtext( "Men"  , side=3, at=0.25, outer=TRUE, line=2 )
+ #mtext( "Women", side=3, at=0.75, outer=TRUE, line=2 )
+ #mtext( "Age at entry", side=3, at=-0.02, outer=TRUE, line=1, adj=0 )
+ mtext( "Time since entry (years)", side=1, outer=TRUE, line=2.5 )
+ mtext( "Probability"             , side=2, outer=TRUE, line=2.5, las=0 )
+ mtext( "Men"  , side=3, at=0.25, outer=TRUE, line=2.5, las=0 )
+ mtext( "Women", side=3, at=0.75, outer=TRUE, line=2.5, las=0 )
+ mtext( "Age"  , side=3, at=1.00, outer=TRUE, line=2.5, las=0, adj=1 )
+ }
> pl10(lp)

> pl10(lpi)
```

4.5 Cumulative risks

From the `pState` objects we can derive the fraction of patients at say 10 and 20 years after baseline that are alive, alive without CVD, who ever had CVD (live or dead), which for the Steno 2 population will illustrate the general effect of the intervention, and for the specific age and sex groups how the intervention may be expected to work in other settings.

There are two main outcomes of which we would like to know: death and CVD; in terms of figure 4.1 "Death" is the probability of being in any of the right hand boxes; "CVD" is the probability of being in any of the lower 6 boxes (regardless of being dead or alive). We can make a table for the Steno population, and a table for the 10 combinations of age and sex we defined.

First the calculations for the Steno2 population — essentially just sums/differences between probabilities in different states:

```
> cr2 <- NArray( list( type = c("Death","CVD"),
+                      when = rownames(pI),
+                       grp = c("Intensive","Conventional","Int-Conv") ) )
> cr2["Death",,"Intensive"   ] <- 1-pI[,4]
> cr2["Death",,"Conventional"] <- 1-pC[,4]
> cr2["CVD"  ,,"Intensive"   ] <- pI[,7]-pI[,1]
> cr2["CVD"  ,,"Conventional"] <- pC[,7]-pC[,1]
> cr2[,,"Int-Conv"] <- cr2[,,"Intensive"] - cr2[,,"Conventional"]
> round( ftable(cr2[,c("10","20"),])*100, 1 )
           grp Intensive Conventional Int-Conv
type  when
Death 10            21.5         30.1     -8.6
      20            49.4         63.6    -14.2
CVD   10            27.5         43.4    -15.8
      20            44.7         62.8    -18.1
```

. . . and then the same for the 10 combinations of age and sex:

```
> cras <- NArray( list( age = seq(45,65,5),
+                       sex = c("M","F"),
+                      type = c("Death","CVD"),
+                      when = dimnames(lp[[1]])[["when"]],
+                       grp = c("Intensive","Conventional","Int-Conv") ) )
> for( ia in dimnames(cras)[["age"]] )
+ for( is in dimnames(cras)[["sex"]] )
+ for( ig in dimnames(cras)[["grp"]][-3] )
+    {
+    cras[ia,is,"Death",,ig] <- 1-lp[[ia,is,ig]][,4]
+    cras[ia,is,"CVD"  ,,ig] <-   lp[[ia,is,ig]][,7]-
+                                 lp[[ia,is,ig]][,1]
+    }
> cras[,,,,3] <- cras[,,,,2] - cras[,,,,1]
> round( ftable( cras[,,,paste(c(10,20))],
+                col.vars=c(3,5),
+                row.vars=c(4,2,1) )*100, 1 )
             type     Death                              CVD
             grp  Intensive Conventional Int-Conv Intensive Conventional Int-Conv
when sex age
10   M   45            6.9         10.6      3.7      20.9         34.0     13.2
         50           11.5         17.5      6.0      24.1         38.9     14.8
         55           18.5         28.2      9.7      26.8         44.7     17.9
         60           30.4         44.9     14.5      32.9         50.0     17.0
         65           46.8         61.2     14.4      37.0         54.5     17.4
     F   45            4.8          6.5      1.7      16.2         26.0      9.8
         50            7.9         11.6      3.7      18.5         32.9     14.4
         55           13.0         18.9      5.9      23.8         37.4     13.6
         60           21.3         30.2      8.9      26.3         42.0     15.7
         65           32.8         47.3     14.5      30.1         47.9     17.9
```

```
20   M   45          21.7        32.6     11.0      39.2        57.7     18.5
         50          33.8        48.9     15.1      43.2        62.4     19.2
         55          49.8        68.0     18.1      45.1        65.5     20.4
         60          69.2        84.6     15.4      50.6        66.8     16.2
         65          84.3        93.5      9.2      50.9        67.0     16.1
     F   45          14.8        20.8      6.0      31.8        47.0     15.2
         50          24.3        35.3     11.0      36.4        55.1     18.7
         55          37.5        52.0     14.5      41.2        59.2     17.9
         60          55.3        70.5     15.2      43.5        61.5     18.0
         65          70.2        86.1     15.9      46.0        63.7     17.7
```

The above table gives the cumulative risks of death and CVD respectively in percent, computed at 10 and 20 years after baseline, but we could of course also show the cumulative risks as curves; first for the Steno 2 population:

```
> tt <- as.numeric( dimnames(cr2)[[2]] )
> par( mfrow=c(1,2), mar=c(2,3,1,1), oma=c(2,0,0,0), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> for( tp in c("Death","CVD") )
+    {
+ matplot( tt, cr2[tp,,1]*100,
+           type="l", lty=1, lwd=4, col="black",
+           ylim=c(0,100), xlab="", yaxs="i", ylab=paste("Cumulative risk of", tp) )
+ matlines( tt, cr2[tp,,2]*100,
+           type="l", lty="11", lwd=4, lend="butt", col="black" )
+    }
> mtext( "Time since baseline (years)", side=1, line=0, outer=TRUE )


> par( mfrow=c(2,2), mar=c(2,3,1,1), oma=c(2,0,0,0), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> for( sx in c("M","F") )
+ for( tp in c("Death","CVD") )
+    {
+ matplot( tt, t( rbind( cras[,sx,tp,,"Intensive"] ) )*100,
+           type="l", lty=1, lwd=1:5, col="black",
+           ylim=c(0,100), xlab="", yaxs="i", ylab=paste("Cumulative risk of",tp) )
+ matlines( tt, t( rbind( cras[,sx,tp,,"Conventional"] ) )*100,
+           type="l", lty=c("AA","55","43","32","22"), lend="butt", lwd=1:5, col="black" )
+    text( 0, 90, sx, font=2, cex=1.5, adj=0 )
+    }
> mtext( "Time since baseline (years)", side=1, line=0, outer=TRUE )


> par( mfrow=c(2,2), mar=c(2,1,1,2), oma=c(1,2,0,0), mgp=c(3,1,0)/1.6, bty="n", las=1 )
> for( sx in c("M","F") )
+ for( tp in c("Death","CVD") )
+    {
+ matplot( tt, t( rbind( cras[,sx,tp,,"Intensive"] ) )*100,
+           type="l", lty=1, lwd=1:5, col="black",
+           xlim=c(0, 20), xaxs="i", xlab="",
+           ylim=c(0,100), yaxs="i", ylab="" )
+ matlines( tt, t( rbind( cras[,sx,tp,,"Conventional"] ) )*100,
+           type="l", lty=c("AA","55","43","32","22"), lend="butt", lwd=1:5, col="black" )
+ axis( side=4, at=0:5*20, labels=if(tp=="Death") NA else 0:5*20 )
+ axis( side=4, at=0:10*10, tcl=-0.4, labels=NA )
+ axis( side=4, at=0:20*5 , tcl=-0.3, labels=NA )
+ axis( side=4, at=0:100  , tcl=-0.2, labels=NA )
+    }
> mtext( "Time since baseline (years)", side=1, line=0, outer=TRUE )
> mtext( c("Women","Cumulative risk (%)","Men"), at=1:3/4, side=2, line=1, outer=TRUE, las=0 )
> mtext( c("Death","CVD"), at=c(1,3)/4, side=3, line=-1, outer=TRUE, las=0 )
```

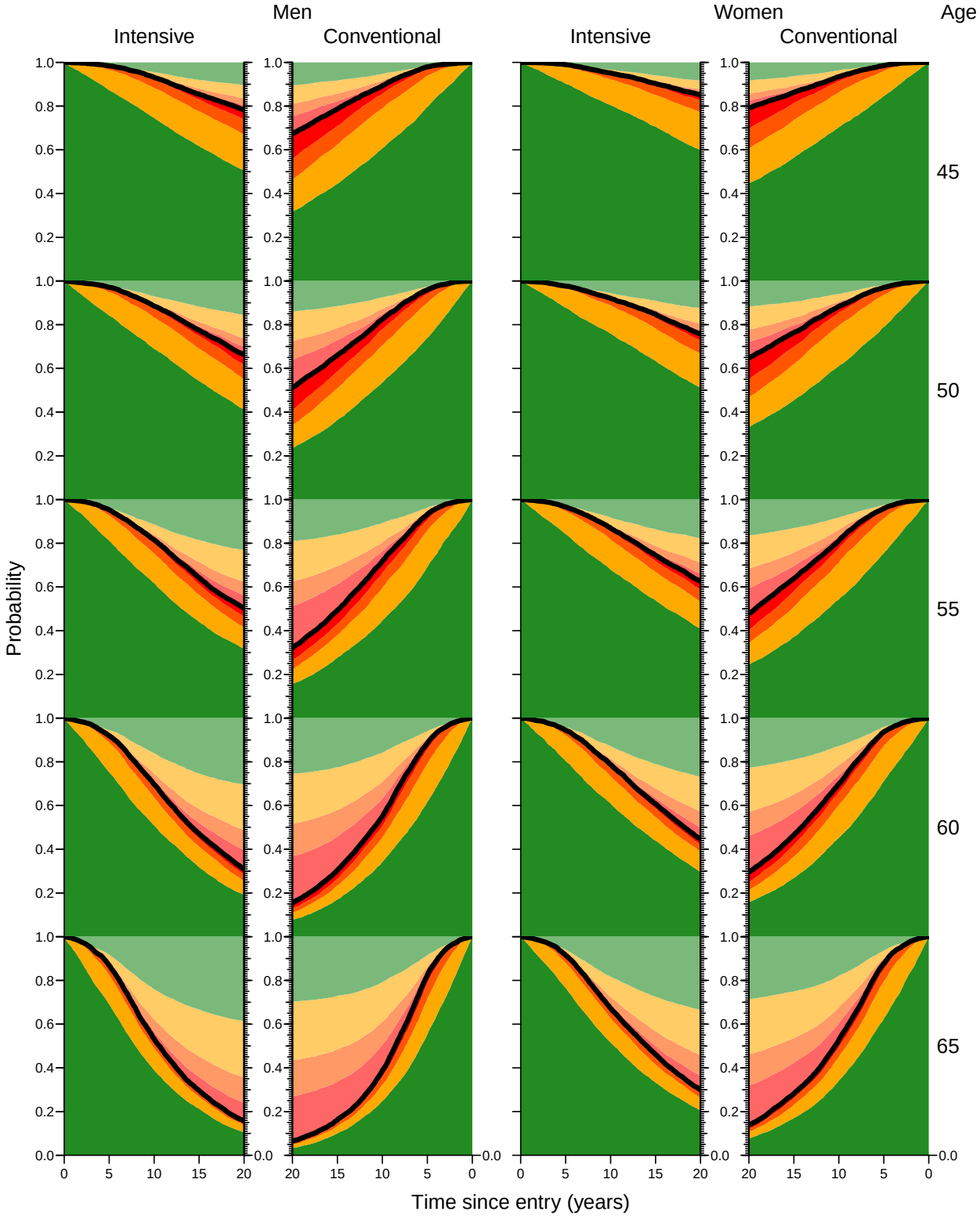

Figure 4.7: *Probabilities of being in different states by age, sex and intervention group, based on models with the same intervention effect across CVD strata. State colouring is the same as in figure 4.1.*

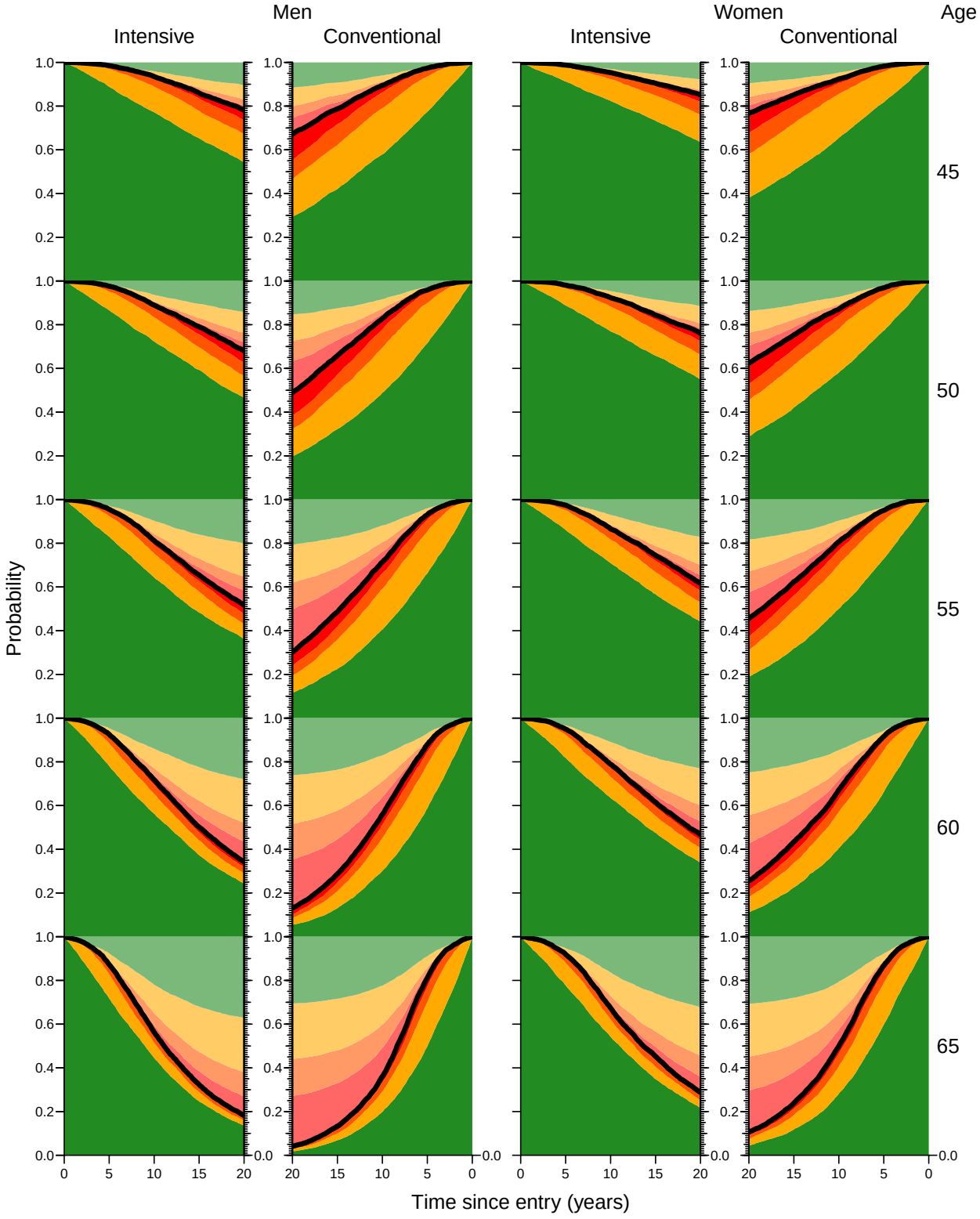

Figure 4.8: *Probabilities of being in different states by age, sex and intervention group, based on models with the differential intervention effect across CVD strata. State colouring is the same as in figure 4.1.*

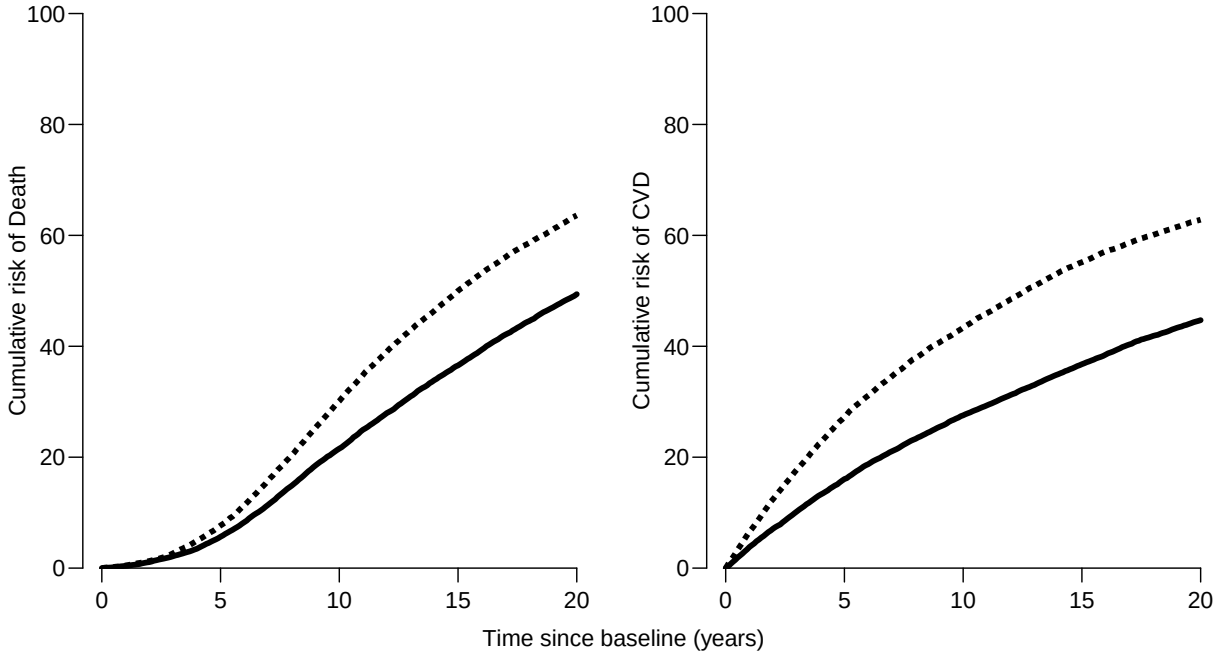

Figure 4.9: *Cumulative risk of death and CVD in the Steno 2 study, as estimated from the multistate model. The full lines is the intensive group; the broken the conventional group.*

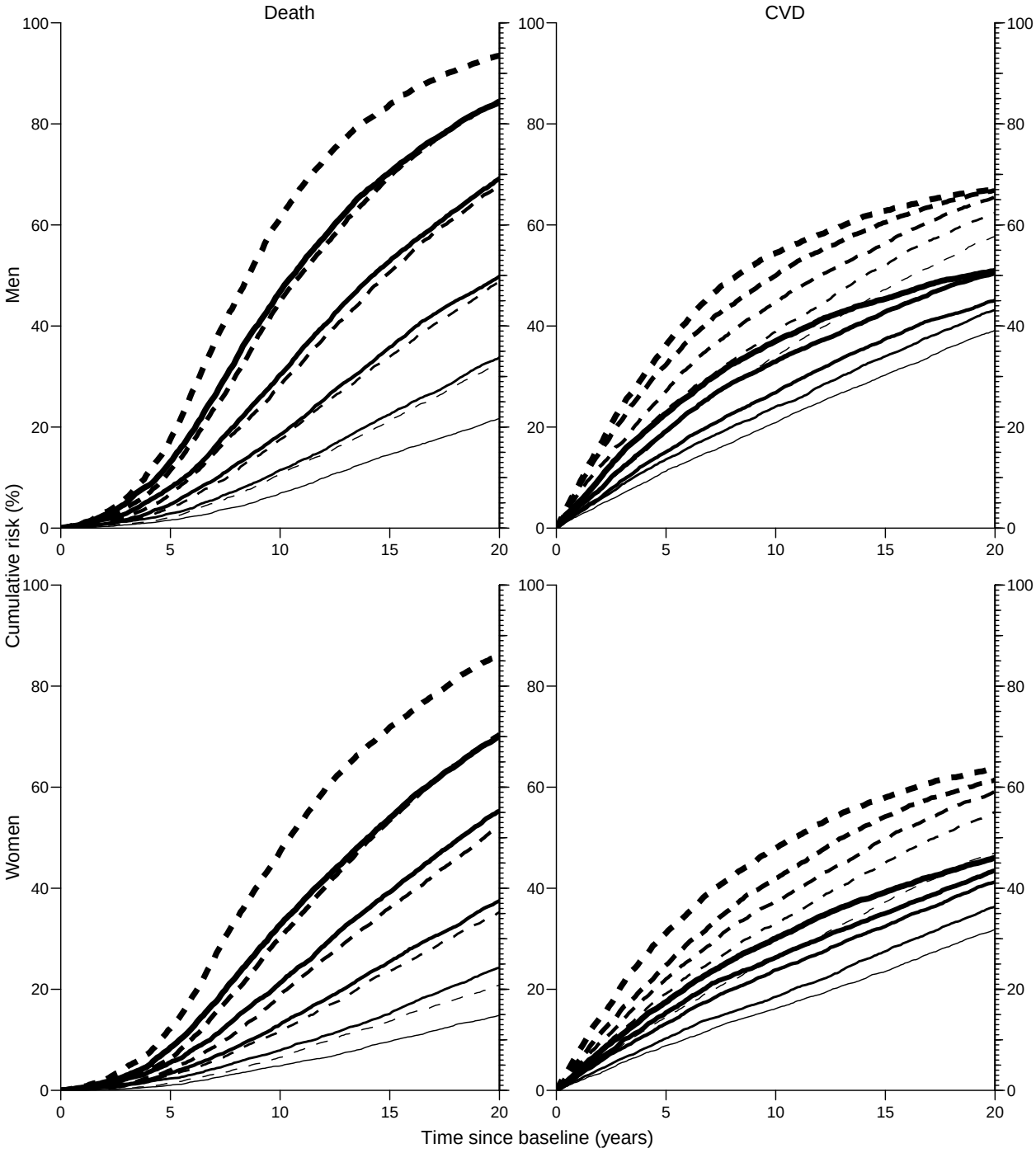

Figure 4.10: *Cumulative risk of death (left) and CVD (right) in the Steno 2 study, as estimated from the multistate model. Top panels are men, bottom panels women; in order of thickness the lines refer to ages 45, 50, 55, 60 and 65 years at baseline. The full lines are for the intensive group; the broken for the conventional group.*

4.6 Years lived with and without CVD

Here we need (for the sake of completeness) both the estimated state-distributions for the Steno2-study (in the file `pSt2.Rda`) and for selected combinations of age and sex (in the file `lp.Rda`):

```
> load( file="../data/pSt2.Rda" )
> load( file="../data/lp.Rda" )
```

We can now compute the sojourn time in each state during the first 20 years after randomization. This is really just a transformation of the matrix of state occupancy probabilities (the `pState` object), namely the integral of the probabilities over time. We devise a small function that does this job; note that the function returns a matrix of the same dimension as the `pState` object; it contains the time spent in each state up to a given time:

```
> cum.area <-
+ function( pp )
+ {
+ if( !inherits(pp,"pState") ) warning( "Not a pState object" )
+ il <- diff( as.numeric(rownames(pp)) )
+ st <- pp * 0
+ pp <- cbind(0,pp) # makes the code simpler
+ for( i in 2:ncol(pp) )
+    st[-1,i-1] <- cumsum( il * ( pp[-1,i  ]-diff(pp[,i  ])/2 -
+                                 ( pp[-1,i-1]-diff(pp[,i-1])/2 ) ) )
+ class( st ) <- "matrix"
+ st
+ }
> cum.area
function( pp )
{
if( !inherits(pp,"pState") ) warning( "Not a pState object" )
il <- diff( as.numeric(rownames(pp)) )
st <- pp * 0
pp <- cbind(0,pp) # makes the code simpler
for( i in 2:ncol(pp) )
   st[-1,i-1] <- cumsum( il * ( pp[-1,i  ]-diff(pp[,i  ])/2 -
                                ( pp[-1,i-1]-diff(pp[,i-1])/2 ) ) )
class( st ) <- "matrix"
st
}
> yI <- cum.area( pI )
> yC <- cum.area( pC )
> lc <- lapply( lp, cum.area )
> lci <- lapply( lpi, cum.area )
> dim( lc ) <- dim( lci ) <- dim( lp )
> dimnames( lc ) <- dimnames( lci ) <- dimnames( lp )
```

With these calculations fixed we can show the years lived with and without CVD in the two groups during the first 20 years of follow-up, and of course also the difference between the groups. First the calculation for the Steno2 population where we get the number of years lived in total, free of CVD, with 1 CVD event (which may not be of much relevance) and with any number of CVD events:

```
> y2 <- NArray( c( list( state = c("Alive","No CVD","1st CVD","Any CVD"),
+                        grp = c(dimnames(lc)[[3]],"Int-Ctr") ) ) )
> y2["No CVD" ,-3] <- c( yI["20","DM"],
+                        yC["20","DM"] )
> y2["1st CVD",-3] <- c( yI["20","1st CVD"],
+                        yI["20","1st CVD"] )
```

```
> y2["Any CVD",-3] <- c( yI["20","1st CVD"]+yI["20","2nd CVD"]+yI["20","3+ CVD"],
+                        yC["20","1st CVD"]+yC["20","2nd CVD"]+yC["20","3+ CVD"] )
> y2["Alive",] <- y2["No CVD",] + y2["Any CVD",]
> y2[,"Int-Ctr"] <- y2[,"Intensive"] - y2[,"Conventional"]
> round( ftable( y2 ), 1 )
        grp Intensive Conventional Int-Ctr
state
Alive            15.6         14.1     1.5
No CVD           12.6         10.0     2.6
1st CVD           2.2          2.2     0.0
Any CVD           3.0          4.1    -1.1
```

Thus we see that the total years gained during the 20 first years of follow-up is 1.6, but 2.6 CVD-free years so the intensive group has actually *fewer* years lived with CVD.

However, these numbers merely tend to illustrate the effects on the specific Steno2 distribution, and they are not generalizable to populations with other age-and sex-distributions. Therefore, we also show results for specific age-groups for each sex.

```
> yl <- NArray( c( dimnames(lc)[-3],
+                  list( state = c("Alive","No CVD","1st CVD","Any CVD"),
+                        grp = c(dimnames(lc)[[3]],"Int-Ctr") ) ) )
> for( ia in dimnames(lc)[[1]] )
+ for( is in dimnames(lc)[[2]] )
+ for( ig in dimnames(lc)[[3]] )
+    {
+    yl[ia,is,"No CVD" ,ig] <- lc[[ia,is,ig]]["20","DM"]
+    yl[ia,is,"1st CVD",ig] <- lc[[ia,is,ig]]["20","1st CVD"]
+    yl[ia,is,"Any CVD",ig] <- lc[[ia,is,ig]]["20","1st CVD"] +
+                              lc[[ia,is,ig]]["20","2nd CVD"] +
+                              lc[[ia,is,ig]]["20","3+ CVD"]
+    }
> yl[,,"Alive",] <- yl[,,"No CVD",] + yl[,,"Any CVD",]
> yl[,,,"Int-Ctr"] <- yl[,,,"Intensive"] - yl[,,,"Conventional"]
> round( ftable( yl[,,-3,], row.vars=c(3,1) ), 1 )
            sex         F                                  M
            grp Intensive Conventional Int-Ctr Intensive Conventional Int-Ctr
state   age
Alive   45           18.9         18.4     0.5      18.3         17.5     0.8
        50           18.1         17.2     0.9      17.4         16.0     1.3
        55           17.0         15.8     1.2      15.8         14.1     1.8
        60           15.4         13.8     1.6      13.8         11.7     2.1
        65           13.5         11.3     2.3      11.4          9.5     1.9
No CVD  45           16.1         14.2     1.9      14.9         12.5     2.4
        50           15.1         12.6     2.6      13.9         11.2     2.7
        55           13.8         11.4     2.4      12.6          9.8     2.8
        60           12.5          9.9     2.5      10.8          8.2     2.6
        65           11.0          8.2     2.8       9.1          6.9     2.2
Any CVD 45            2.8          4.3    -1.4       3.4          5.0    -1.6
        50            3.0          4.7    -1.7       3.5          4.8    -1.4
        55            3.2          4.4    -1.2       3.2          4.2    -1.0
        60            2.9          3.8    -0.9       2.9          3.5    -0.5
        65            2.5          3.1    -0.6       2.3          2.6    -0.3
```

Thus we see that the extra years gained by the intensive during the first 20 years is between 0.5 and 2.8; largest among men and increasing by age at entry to the study. The years lived without any CVD is about 2.5 years, largely the same in all age-classes and both for men and women. This means that the years alive with CVD is largest in the conventional group about 1.5 years larger for persons entering at age 45 decreasing to about 0.5 years for persons entering in age 65.

Thus, the major treatment results seem to be about 2 more years of life, and 2.5 years longer without CVD during the first 20 years of follow-up.

Chapter 5

Microvascular events

The microvascular outcomes are measured at 6 different visits, and changes of status are therefore inherently *interval censored*. We will impute exact transition times using a random uniformly distributed time between the two visits; and we shall look into the effect of the imputation by assessing the difference between resulst from different imputaions.

5.1 Useful functions for interval censoring

To this end we have devised functions to perform some of the tasks:

```
> source( "usefuns.R" )
```

The function `rpart` generates a vector with `a` as the first element, `b` as the last and `n` uniformly distributed points between them. This is intended for generating the transition times trough (possibly more than one) intermediate step, when only the status at visit times `a` and `b` are known:

```
> rpart
function(n,a,b) a+(b-a)*c(0,sort(runif(n)),1)
> rpart( 2, 3, 7 )
[1] 3.00000 4.30165 5.76567 7.00000
> rpart( 2, 2, 7 )
[1] 2.000000 6.646254 6.937107 7.000000
```

The function `lx.gen` generates a data frame with the skeleton of a `Lexis` object "by hand", imputing all intermediate states and transition times:

```
> lx.gen
function( st, dt, id, incr=TRUE )
{
# Generate Lexis records for one person in states st at dates dt.
#
# st is a vector of states (possibly with missing)  - rows of rt
# dt is a vector dates of visits                    - rows of rd
# id is a number, used for labeling (as lex.id)     - rownames of these
# incr indicates if only increasing states should be allowed - meaning
# that all decreases in states are overridden by the previous larger,
# effectively redefining state to "highest state achieved thus far".
# res.df is the resulting dataframe (Lexis object) with follow-up of
# ONE person
res.df <- NULL
# only states and dates known
```

```
dt <- dt[!is.na(st)]
st <- st[!is.na(st)]
# number of points with known status
np <- length( st )
# only ever output anything substantial if more than one obs
if( np > 1 )
{
# only assume increasing states
if( incr ) for( i in 2:np ) st[i] <- max( st[i-0:1] )
# loop over states seen
for( i in 2:np )
   {
   if( st[i]==st[i-1] )
     {
     lex.Cst <- lex.Xst <- st[i]
     lex.dur <- dt[i] - dt[i-1]
     per <- dt[i-1]
     res.df <- rbind( res.df,
                      data.frame( per=per, lex.dur, lex.Cst, lex.Xst, lex.id=id ) )
     }
   if( st[i]!=st[i-1] )
     {
     # no. of transitions seen
     nt <- abs(st[i]-st[i-1])
     tm <- rpart( nt, dt[i-1], dt[i] )
     lex.Cst <- st[i-1]:st[i]
     lex.Xst <- c(lex.Cst[-1],rev(lex.Cst)[1])
     lex.dur <- diff( tm )
     per <- tm[-length(tm)]
     res.df <- rbind( res.df,
                      data.frame( per=per, lex.dur, lex.Cst, lex.Xst, lex.id=id ) )
     }
   }
} else res.df <- data.frame( per = dt,   # The case of ony one visit
                          lex.dur = 1/12, # risk is set to 1 month
                          lex.Cst = st,
                          lex.Xst = st,
                           lex.id = id )
res.df
}
```

The function `make.dlex` merges a "raw" object of transition times (generated by `lx.gen`) with dates of death and other relevant dates and other variables from the baseline visit in the data frame `mic`.

```
> make.dlex
function( rt, rd, incr=TRUE )
{
lx.new <- NULL
for( i in 1:nrow(rt) )
   lx.new <- rbind( lx.new,
                    lx.gen( rt[i,],
                            rd[i,],
                       id = as.numeric(rownames(rt)[i]),
                     incr = incr ) )
# Get the exit time and the detah status
dd <- mic[,c("ID","doDth","doEnd")]
dd <- dd[match(unique(dd$ID),dd$ID),]
# The last record in the Lexis object of each person
ll <- merge( subset( lx.new, per==ave( per, lex.id, FUN=max ) ), dd,
             by.x="lex.id", by.y="ID", all.x=TRUE )
# Change this record to represent follow-up from last visit to death in last known state
ll$lex.Cst <- ll$lex.Xst
ll$per     <- ll$per+ll$lex.dur
ll$lex.Xst <- ifelse( is.na(ll$doDth),
```

```
                      ll$lex.Cst,
                      paste("Dead(",ll$lex.Cst,")",sep="") )
ll$lex.dur <- pmin( ll$doDth, ll$doEnd, na.rm=TRUE ) - ll$per
ll <- subset( ll, lex.dur>0 )
lx.new <- rbind( lx.new, ll[,names(lx.new)] )
lx.new <- merge( lx.new,
                 mic[mic$visit==1,c("ID","sex","doBth","doVis","doBlind","doEnd","doDth","allocation
                 by.x ="lex.id", by.y="ID", all.x=TRUE )
lx.new <- transform( lx.new,
                     age = per-doBth,
                     tfe = per-doVis )
attr( lx.new, "class" ) <- c("Lexis","data.frame")
attr( lx.new, "time.scales" ) <- c("per","age","tfe")
attr( lx.new, "time.since"  ) <- c("","","")
ol <- with( lx.new, order(lex.id,per) )
Relevel( lx.new[ol,] )
}
```

With these functions in place we can now analyze occurrence of the three types of microvascular events.

5.2 Retinopathy

The analysis will be done for the maximal score of left and right eye, that is maximum of `eds_l` and `eds_r`.

5.2.1 Data overview

First we provide an overview of the actually occurring retinopathy patterns in the data:

```
> library( Epi )
> library( survival )
> clear()
> load( file="../data/mic.Rda" )
> names( mic )
 [1] "ID"         "allocation" "sex"        "doBth"      "doVis"      "doESRD"
 [7] "doEye"      "doPhC"      "doBlind"    "doEnd"      "doDth"      "visit"
[13] "aer_level"  "med_aer"    "eds_l"      "eds_r"      "autoprog"   "periprog"
> with( mic, ftable(visit,!is.na(eds_r),!is.na(eds_l),useNA="ifany",col.vars=c(1,3)) )
      visit     1          2          3          4          5          6
            FALSE TRUE FALSE TRUE FALSE TRUE FALSE TRUE FALSE TRUE FALSE TRUE

FALSE           0    3    92    1     0    1     0    1    14    1    62    0
TRUE            2  155     4   59     1  147     1  126     0   78     0    0
> with( mic, table( eds_l, eds_r) )
                          eds_r
eds_l                      No retinopathy Minimal non-proliferative
  No retinopathy                      248                        35
  Minimal non-proliferative            39                        96
  Moderate nonproliferative             6                        14
  Pre-proliferative                     0                         0
  Proliferative                         0                         0
  Photocoagulation                      1                         3
                          eds_r
eds_l                      Moderate nonproliferative Pre-proliferative Proliferative
  No retinopathy                                   3                0             0
  Minimal non-proliferative                        8                1             3
  Moderate nonproliferative                       28                2             1
  Pre-proliferative                                1                7             2
  Proliferative                                    2                2            11
```

```
  Photocoagulation                                4                1             5
                          eds_r
eds_l                      Photocoagulation
  No retinopathy                          1
  Minimal non-proliferative               0
  Moderate nonproliferative               5
  Pre-proliferative                       1
  Proliferative                           5
  Photocoagulation                       30
> with( mic, plot( jitter(as.integer(eds_r),2),
+                  jitter(as.integer(eds_l),2), pch=16, cex=0.8 ) )
```

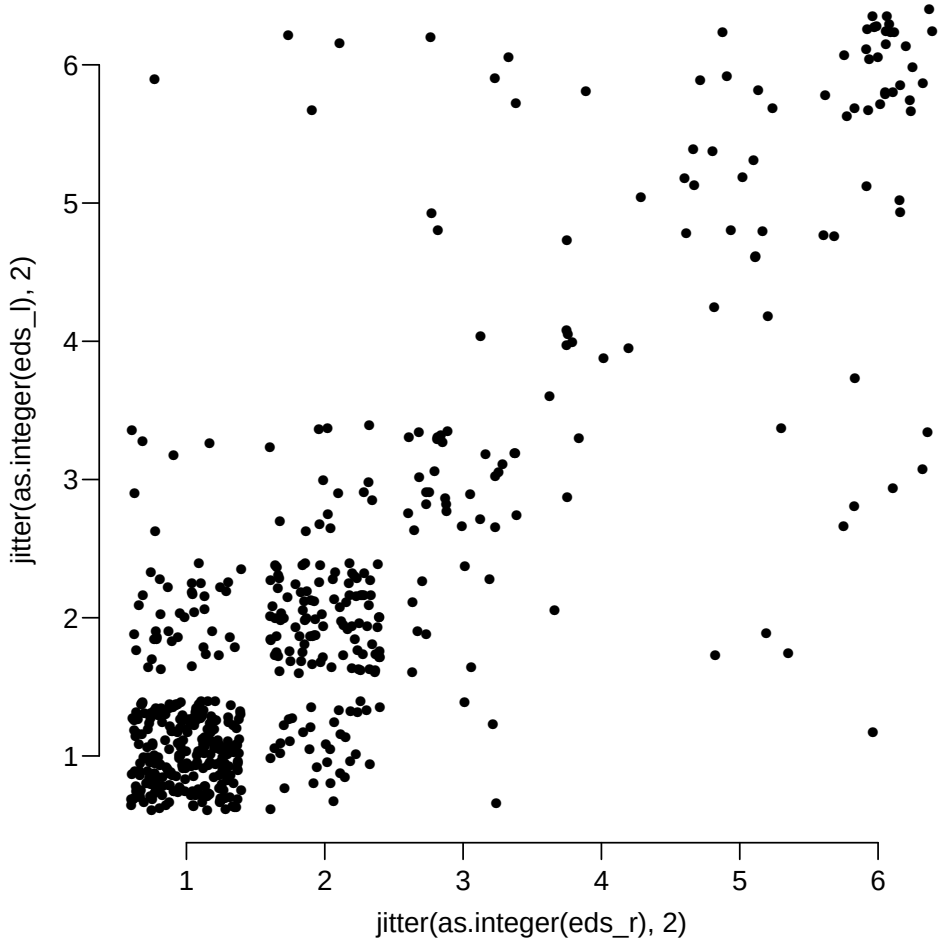

Figure 5.1: *Relationship between left and right eds-score, jittered*

For the sake of consistency we form a variable which is the maximum of the two sides.

```
> mic <- transform( mic, eds = pmax(as.integer(eds_r),
+                                   as.integer(eds_l),na.rm=TRUE) )
> with( mic, table( eds ) )
eds
  1   2   3   4   5   6
259 170  59  11  22  59
```

We then generate the two 160×6 matrices of states (`rt`, maximal retinopathy score) and dates (`rd`):

```
> rd <- with( mic, tapply( doVis, list(ID,visit), mean ) )
> rt <- with( mic, tapply(   eds, list(ID,visit), mean ) )
> cbind( rt, rd )[1:10,]
   1  2  3  4  5  6        1        2        3        4        5        6
1  1 NA  1  1  1 NA 1993.357 1995.158 1997.135 2001.097 2006.712 2014.802
2  2 NA  2  6  6 NA 1993.357 1995.377 1997.781 2001.814 2006.868 2014.589
3  2 NA  2  4 NA NA 1993.362 1995.180 1997.151 2001.102       NA       NA
4  1  1  1  1 NA NA 1993.362 1995.120 1997.110 2001.066       NA       NA
5  1 NA  1 NA NA NA 1993.162 1995.002 1997.707       NA       NA       NA
6  1 NA  3  3  6 NA 1993.201 1995.008 1997.036 2001.025 2006.736 2015.029
7  1  2  3  6 NA NA 1993.203 1995.043 1997.077 2001.058       NA       NA
8  2  2  2  2  2 NA 1993.209 1995.829 1997.729 2001.655 2006.903       NA
9  2  1 NA NA NA NA 1993.209 1995.021       NA       NA       NA       NA
10 1  2  2  6 NA NA 1993.075 1995.024 1997.020 2001.023 2006.682       NA
> table( rt[,1] )
  1   2   3   4   5   6
108  27  10   5   6   4
```

This gives the opportunity to generate the actually observed patterns of status:

```
> rr <- gsub("NA","-",apply( rt, 1, paste, collapse=""))
> tt <- table( rr )
> rn <- gsub("-","7",rownames( tt ))
> addmargins( tt[order(rn)] )
rr
11111- 11112- 1111-- 11121- 11122- 11123- 11126- 1112-- 11131- 111--- 11221- 11222-
     4      3      4      1      1      1      1      2      1      1      1      1
11223- 1122-- 1123-- 11266- 112--- 113--- 11-1-- 11---- 12222- 12223- 12224- 12226-
     1      1      1      1      2      1      2      1      1      1      1      2
1223-- 1226-- 122--- 1236-- 1322-- 133--- 13---- 1-111- 1-112- 1-11-- 1-122- 1-126-
     1      1      2      1      2      2      1     14      2      6      3      2
1-12-- 1-132- 1-136- 1-166- 1-1--- 1-212- 1-222- 1-22-- 1-232- 1-233- 1-23-- 1-26--
     2      2      1      1      2      1      4      3      1      1      2      1
1-2--- 1-323- 1-332- 1-336- 1-33-- 1-66-- 1--5-- 1----- 2112-- 21222- 2122-- 21----
     3      1      1      1      2      1      1      4      1      1      1      1
22112- 22222- 2222-- 222-6- 222--- 223--- 233--- 2-21-- 2-222- 2-225- 2-232- 2-236-
     1      2      1      1      1      1      1      1      1      1      1      2
2-23-- 2-24-- 2-266- 2-26-- 2-2--- 2-366- 2-3--- 2-6-6- 31322- 33---- 3-22-- 3-322-
     1      1      2      1      1      1      1      1      1      1      1      1
3-33-- 3-355- 3-466- 3-46-- 3-566- 3-66-- 43566- 45666- 4-466- 4-46-- 4-666- 555---
     1      1      1      1      1      1      1      1      1      1      1      2
5-55-- 5-65-- 5-66-- 6-66-- 6-6---    Sum
     2      1      1      2      2    160
```

5.2.2 Generating transitions

We again load the functions to generate the Lexis objects of transitions between the states:

```
> source( "usefuns.R" )
```

We now simulate exact transition times to generate data sets and inspect the difference between them:

```
> ret1 <- make.dlex( rt, rd, incr=TRUE )
> ret2 <- make.dlex( rt, rd, incr=FALSE )
> summary( ret1 )
Transitions:
     To
From    1   2   3  4  5   6 Dead(1) Dead(2) Dead(3) Dead(4) Dead(5) Dead(6)  Records:
  1   159  70   0  0  0   0      21       0       0       0       0       0       250
  2     0 156  52  0  0   0       0      27       0       0       0       0       235
```

```
  3     0   0  78 30  0   0       0       0      20       0       0       0       128
  4     0   0   0  8 33   0       0       0       0       1       0       0        42
  5     0   0   0  0 17  32       0       0       0       0       5       0        54
  6     0   0   0  0  0  73       0       0       0       0       0      19        92
  Sum 159 226 130 38 50 105      21      27      20       1       5      19       801

Transitions:
     To
From   Events: Risk time:  Persons:
  1          91     805.02       108
  2          79     702.75        97
  3          50     377.03        62
  4          34      68.01        35
  5          37     106.81        39
  6          19     357.99        36
  Sum       310    2417.60       160
> summary( ret2 )
Transitions:
     To
From    1   2   3  4  5   6 Dead(1) Dead(2) Dead(3) Dead(4) Dead(5) Dead(6)  Records:
  1   175  76   0  0  0   0       23       0       0       0       0       0       274
  2    11 165  54  0  0   0        0      29       0       0       0       0       259
  3     0  13  54 31  0   0        0       0      16       0       0       0       114
  4     0   0   1  7 33   0        0       0       0       1       0       0        42
  5     0   0   0  0 18  32        0       0       0       0       6       0        56
  6     0   0   0  0  1  72        0       0       0       0       0      18        91
  Sum 186 254 109 38 52 104       23      29      16       1       6      18       836

Transitions:
     To
From   Events: Risk time:  Persons:
  1          99     878.97       115
  2          94     744.95       100
  3          60     258.27        63
  4          35      67.48        35
  5          38     110.31        39
  6          19     357.61        36
  Sum       345    2417.60       160
```

Note that what differs between the data sets is the transitions and the relative distribution of the risk time between states; the total risk time is fixed for each person. The two datasets differ in transitions because *improvement* (decreasing state number) is allowed in the second.

5.2.3 Reversible retinopathy

As there are 40 observed transitions to *better* status we shall use the extended definition of transition that allows this too, except that transition from "Photocoagulation" to "Proliferative" is not allowed.

Now we set the seed explicitly so that results are reproducible

```
> set.seed( 4667571 )
> ret <- make.dlex( rt, rd, incr=FALSE )
> summary( ret )
Transitions:
     To
From    1   2   3  4  5   6 Dead(1) Dead(2) Dead(3) Dead(4) Dead(5) Dead(6)  Records:
  1   175  76   0  0  0   0       23       0       0       0       0       0       274
  2    11 165  54  0  0   0        0      29       0       0       0       0       259
  3     0  13  54 31  0   0        0       0      16       0       0       0       114
  4     0   0   1  7 33   0        0       0       0       1       0       0        42
```

```
  5     0   0   0  0 18  32       0       0       0       0       6       0        56
  6     0   0   0  0  1  72       0       0       0       0       0      18        91
  Sum 186 254 109 38 52 104      23      29      16       1       6      18       836

Transitions:
     To
From   Events: Risk time:  Persons:
  1         99     873.08       115
  2         94     757.05       100
  3         60     250.40        63
  4         35      56.00        35
  5         38     110.47        39
  6         19     370.59        36
  Sum      345    2417.60       160
```

```
> levels( ret )
 [1] "1"       "2"       "3"       "4"       "5"       "6"       "Dead(1)" "Dead(2)"
 [9] "Dead(3)" "Dead(4)" "Dead(5)" "Dead(6)"
> levels( mic$eds_l )
[1] "No retinopathy"            "Minimal non-proliferative" "Moderate nonproliferative"
[4] "Pre-proliferative"         "Proliferative"             "Photocoagulation"
> newl <- c("None","Min","Mod","PrePr","Pro","PhC")
> levels( ret$lex.Cst ) <-
+ levels( ret$lex.Xst ) <- c( newl, paste("D(",newl,")",sep="") )
> levels( ret )
 [1] "None"     "Min"      "Mod"      "PrePr"    "Pro"      "PhC"      "D(None)"
 [8] "D(Min)"   "D(Mod)"   "D(PrePr)" "D(Pro)"   "D(PhC)"
```

There is one illegal transition in the observed data; namely from photocoagulation to proliferative:

```
> subset( ret, lex.Cst=="PhC" & lex.Xst=="Pro" )
    lex.id      per   lex.dur lex.Cst lex.Xst  sex   doBth    doVis doBlind    doEnd
598    120 1997.748 0.9464413     PhC     Pro Male 1939.29 1993.775      NA 2005.179
       doDth   allocation      age      tfe
598 2005.179 Conventional 58.45859 3.972621
> subset( ret, lex.id==120 )
    lex.id      per   lex.dur lex.Cst lex.Xst  sex   doBth    doVis doBlind    doEnd
600    120 1993.775 0.8496191     Pro     PhC Male 1939.29 1993.775      NA 2005.179
597    120 1994.625 3.1230024     PhC     PhC Male 1939.29 1993.775      NA 2005.179
598    120 1997.748 0.9464413     PhC     Pro Male 1939.29 1993.775      NA 2005.179
599    120 1998.695 3.0042773     Pro     Pro Male 1939.29 1993.775      NA 2005.179
601    120 2001.699 3.4798084     Pro  D(Pro) Male 1939.29 1993.775      NA 2005.179
       doDth   allocation      age       tfe
600 2005.179 Conventional 54.48597 0.0000000
597 2005.179 Conventional 55.33559 0.8496191
598 2005.179 Conventional 58.45859 3.9726215
599 2005.179 Conventional 59.40503 4.9190628
601 2005.179 Conventional 62.40931 7.9233402
> ret[ret$lex.id==120 & ret$per>1995,"lex.Cst"] <- "PhC"
> ret[ret$lex.id==120,"lex.Xst"] <- gsub("Pro","PhC",ret[ret$lex.id==120,"lex.Xst"] )
> subset( ret, lex.id==120 )
    lex.id      per   lex.dur lex.Cst lex.Xst  sex   doBth    doVis doBlind    doEnd
600    120 1993.775 0.8496191     Pro     PhC Male 1939.29 1993.775      NA 2005.179
597    120 1994.625 3.1230024     PhC     PhC Male 1939.29 1993.775      NA 2005.179
598    120 1997.748 0.9464413     PhC     PhC Male 1939.29 1993.775      NA 2005.179
599    120 1998.695 3.0042773     PhC     PhC Male 1939.29 1993.775      NA 2005.179
601    120 2001.699 3.4798084     PhC  D(PhC) Male 1939.29 1993.775      NA 2005.179
       doDth   allocation      age       tfe
600 2005.179 Conventional 54.48597 0.0000000
597 2005.179 Conventional 55.33559 0.8496191
598 2005.179 Conventional 58.45859 3.9726215
599 2005.179 Conventional 59.40503 4.9190628
601 2005.179 Conventional 62.40931 7.9233402
```

With this in place we can now show the transitions through retinopathy states in the two groups; first defining the colors to use for the different states:

```
> clr <- c("forestgreen",heat.colors(7)[5:1])
> clr <- c(clr,rgb(t(col2rgb(clr)*0.6+255*0.4),max=255))
> apos = rep(0.45,12)
> apos[c(3,6,9)] <- 0.25
> par( mfrow=c(1,2), oma=c(0,0,0,0) )
> for( lv in levels(ret$allocation) )
+    {
+ boxes( subset( ret, allocation==lv ),
+        boxpos=list(x=rep(c(20,80),each=6)+7-14*(lv=="Intensive"),
+                    y=rep(seq(90,5,,6),2)),
+        show.BE="noz", scale.R=100, pos.arr=apos,
+        col.bg=clr,
+        col.border=c(clr[1:6],rep("black",6)),
+        col.txt=rep(c("white","black"),each=6) )
+ text(10+5-10*(lv=="Intensive"),98,lv,cex=1.5,font=2,adj=0)
+ apos[9] <- apos[1]
+    }
```

From figure 5.2 it is clear that the risk time in the states "Mod", "PrePr" and "Pro" is quite small, indicating that persons travel quite quicly through these, on to "PhC" — apparent from the transition rates.

5.2.3.1 Analysis of rates

We now analyze transition rates between the two groups; what we do is to assume different transition intensities between levels of retinopathy, and fit models where intervention effect is:

- different between levels
- changing linearly across levels — separately for up/down transitions; only relevant and sensible for mortality rates
- is constant across levels

We fit three different sets of models; one set for `i`ncreasing retinopathy status (deterioration) and one for `d`ecreasing (improvement), and a set of models for mortality:

In order to get readable code we use the stacked follow-up data for analysis. We create subsets of the stacked data for analysis of up-, resp. down- transitions, and shrink the transition levels to those actually present in the subset:

```
> stret <- stack( ret )
> ( trs <- levels( stret$lex.Tr ) )
 [1] "None->Min"       "None->D(None)"   "Min->None"       "Min->Mod"
 [5] "Min->D(Min)"     "Mod->Min"        "Mod->PrePr"      "Mod->D(Mod)"
 [9] "PrePr->Mod"      "PrePr->Pro"      "PrePr->D(PrePr)" "Pro->PhC"
[13] "Pro->D(Pro)"     "PhC->D(PhC)"
> ( uptr <- trs[c(1,4,7,10,12)] )
[1] "None->Min"  "Min->Mod"   "Mod->PrePr" "PrePr->Pro" "Pro->PhC"
> ( dntr <- trs[c(9,6,3)] )
[1] "PrePr->Mod" "Mod->Min"   "Min->None"
> ( ddtr <- trs[grep("D",trs)] )
[1] "None->D(None)"   "Min->D(Min)"     "Mod->D(Mod)"     "PrePr->D(PrePr)"
[5] "Pro->D(Pro)"     "PhC->D(PhC)"
```

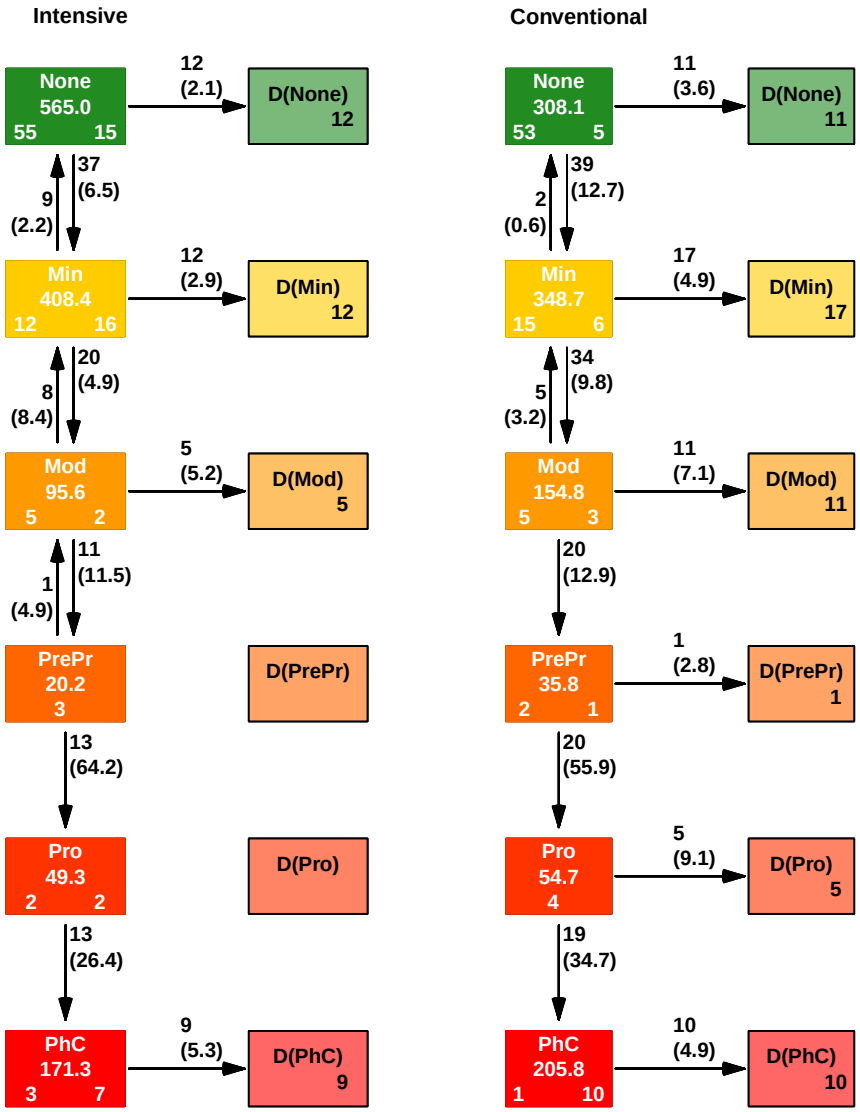

Figure 5.2: *Transitions between states of retinopathy and death, allowing improvements in retinopathy status (except from Photocoagulation state). The number in the center of the boxes is the person-years, the numbers at the bottom is the number of patients starting, resp. ending their follow-up in each box, and the numbers on the arrows are the number of transitions and the overall transition rates per 100 PY.*

```
> upst <- subset( stret, lex.Tr %in% uptr )
> dnst <- subset( stret, lex.Tr %in% dntr )
> ddst <- subset( stret, lex.Tr %in% ddtr )
> upst$lex.Tr <- factor( upst$lex.Tr )
> dnst$lex.Tr <- factor( dnst$lex.Tr )
```

```
> up.sep <- coxph( Surv( tfe, tfe+lex.dur, lex.Fail ) ~
+                  strata(lex.Cst) + allocation:lex.Tr,
+                  data = upst )
> up.con <- update( up.sep, . ~ . - allocation:lex.Tr + allocation )
```

```
> dn.sep <- update( up.sep, data = dnst )
> dn.con <- update( dn.sep, . ~ . - allocation:lex.Tr + allocation )
> dd.sep <- update( up.sep, data = ddst )
> dd.lin <- update( dd.sep, . ~ . - allocation:lex.Tr + allocation + allocation:as.integer(lex.Cst)
> dd.con <- update( dd.sep, . ~ . - allocation:lex.Tr + allocation )
```

For the purpose of prediction it is a bit more convenient to fit the models to the original Lexis object. We need two help-variables, namely the indicators of increase, resp. decrease in retinopathy score:

```
> ( llev <- levels(ret)[1:6] )
[1] "None"  "Min"   "Mod"   "PrePr" "Pro"   "PhC"
> ( dlev <- levels(ret)[7:12] )
[1] "D(None)"  "D(Min)"   "D(Mod)"   "D(PrePr)" "D(Pro)"   "D(PhC)"
> ret <- transform( ret, tr.up = as.integer(lex.Cst)<as.integer(lex.Xst) &
+                               lex.Xst %in% llev,
+                       tr.dn = as.integer(lex.Cst)>as.integer(lex.Xst) &
+                               lex.Xst %in% llev )
> ftable( xtabs( cbind(tr.up,tr.dn) ~ lex.Cst + lex.Xst, data=ret )
+         [1:6,1:6,], row.vars=c(3,1) )
              lex.Xst None Min Mod PrePr Pro PhC
      lex.Cst
tr.up None               0  76   0     0   0   0
      Min                0   0  54     0   0   0
      Mod                0   0   0    31   0   0
      PrePr              0   0   0     0  33   0
      Pro                0   0   0     0   0  32
      PhC                0   0   0     0   0   0
tr.dn None               0   0   0     0   0   0
      Min               11   0   0     0   0   0
      Mod                0  13   0     0   0   0
      PrePr              0   0   1     0   0   0
      Pro                0   0   0     0   0   0
      PhC                0   0   0     0   0   0
> u.sep <- coxph( Surv( tfe, tfe+lex.dur, tr.up ) ~
+                 strata(lex.Cst) + lex.Cst:allocation,
+                 data = subset(ret,lex.Cst %in% llev[1:5]) )
> u.com <- update( u.sep, . ~ . - lex.Cst:allocation + allocation )
> d.sep <- coxph( Surv( tfe, tfe+lex.dur, tr.dn ) ~
+                 strata(lex.Cst) + lex.Cst:allocation,
+                 data = subset(ret,lex.Cst %in% llev[2:6]) )
> d.com <- update( d.sep, . ~ . - lex.Cst:allocation + allocation )
> D.sep <- coxph( Surv( tfe, tfe+lex.dur, lex.Xst %in% dlev ) ~
+                 strata(lex.Cst) + lex.Cst:allocation,
+                 data = ret )
> D.lin <- update( D.sep, . ~ . - lex.Cst:allocation + allocation + allocation:as.integer(lex.Cst) )
> D.com <- update( D.sep, . ~ . - lex.Cst:allocation + allocation )
```

We can check that we get the same by the two approaches:

```
> round( cbind( ci.exp(u.sep), ci.exp(up.sep) ), 3 )
                                 exp(Est.)  2.5% 97.5% exp(Est.)  2.5% 97.5%
lex.CstNone:allocationIntensive      0.661 0.420 1.041     0.661 0.420 1.041
lex.CstMin:allocationIntensive       0.732 0.420 1.276     0.732 0.420 1.276
lex.CstMod:allocationIntensive       0.888 0.421 1.874     0.888 0.421 1.874
lex.CstPrePr:allocationIntensive     0.747 0.348 1.604     0.747 0.348 1.604
lex.CstPro:allocationIntensive       0.847 0.398 1.804     0.847 0.398 1.804
> round( cbind( ci.exp(d.sep), ci.exp(dn.sep) ), 3 )
                                    exp(Est.)  2.5%  97.5%    exp(Est.)  2.5%  97.5%
lex.CstMin:allocationIntensive          5.178 1.109 24.168        5.178 1.109 24.168
lex.CstMod:allocationIntensive          3.085 0.980  9.714        3.085 0.980  9.714
lex.CstPrePr:allocationIntensive 29588452.321 0.000    Inf 29588452.300 0.000    Inf
```

```
> round( cbind( ci.exp(D.sep), ci.exp(dd.sep) ), 3 )
                                exp(Est.)  2.5% 97.5% exp(Est.)  2.5% 97.5%
lex.CstNone:allocationIntensive     0.474 0.206 1.088     0.474 0.206 1.088
lex.CstMin:allocationIntensive      0.480 0.224 1.029     0.480 0.224 1.029
lex.CstMod:allocationIntensive      0.768 0.264 2.233     0.768 0.264 2.233
lex.CstPro:allocationIntensive      0.000 0.000   Inf     0.000 0.000   Inf
lex.CstPhC:allocationIntensive      1.133 0.459 2.798     1.133 0.459 2.798
```

We use a Wald test to check whether the effect of intervention is the same across categories:

```
> ( CC <- t(matrix( c(1,-1,rep(0,5)), 6, 5 ) ) )
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1   -1    0    0    0    0
[2,]    0    1   -1    0    0    0
[3,]    0    0    1   -1    0    0
[4,]    0    0    0    1   -1    0
[5,]    0    0    0    0    1   -1

> Wald( u.sep, ctr.mat=CC[1:4,1:5] )

    Chisq      d.f.         P
0.5889057 4.0000000 0.9642885

> Wald( d.sep, ctr.mat=CC[1:2,1:3] )

    Chisq      d.f.         P
0.2792115 2.0000000 0.8697010

> Wald( D.sep, ctr.mat=CC[1:4,1:5] )

    Chisq      d.f.         P
2.6885343 4.0000000 0.6112224

> round( ci.exp(D.lin,pval=TRUE), 3 )
                                          exp(Est.)  2.5% 97.5%     P
allocationConventional                        2.491 1.132 5.480 0.023
allocationIntensive:as.integer(lex.Cst)       1.131 0.901 1.421 0.289
```

We see that there is no evidence that the HR between intensive and conventional differ across the states for any of the three types of transitions; thus we assume that it is constant across retinopathy states. Thus based on the Cox-models we can report the three HRs associated with being in the intensive group:

```
> HR <- rbind( ci.exp(u.com,pval=TRUE),
+              ci.exp(d.com,pval=TRUE),
+              ci.exp(D.com,pval=TRUE) )
> rownames( HR ) <- c("Increasing ret","Decreasing ret","Death")
> HR <- cbind( 1/HR[,c(1,3,2)], HR[,c(4,1:3)] )
> colnames(HR)[c(1,5)] <- c("Int vs Conv","Conv vs Int")
> round( HR, 3 )
              Int vs Conv 97.5%  2.5%     P Conv vs Int  2.5% 97.5%
Increasing ret       0.739 0.561 0.973 0.031       1.353 1.028 1.781
Decreasing ret       3.932 1.607 9.620 0.003       0.254 0.104 0.622
Death                0.572 0.370 0.883 0.012       1.748 1.132 2.701
```

From the Cox-models we see that there is almost a 25% smaller HR of increasing retinopathy status in the intensive group (HR 0.76; 0.58–1.01), a significantly *larger* HR of improving the retinopathy status (HR 3.4; 1.4–8.2) and a significantly lower mortality when controlling for retinopathy status (HR 0.49; 0.31–0.79), the latter pretty much confirming the overall analyses of mortality.

5.2.3.2 Parametric modeling

In order to get a more credible model for the retinopathy changes and the mortality, we fit a corresponding set of Poisson models where we also take age and sex into account. To this end we split the follow-up data in 1-month intervals:

```
> timeScales( ret )
[1] "per" "age" "tfe"
> sret <- splitLexis( ret, seq(0,24,1/12), "tfe" )
> summary( sret )
Transitions:
     To
From     None  Min  Mod PrePr  Pro  PhC D(None) D(Min) D(Mod) D(PrePr) D(Pro) D(PhC)
  None  10590   76    0     0    0    0      23      0      0         0      0      0
  Min      11 9240   54     0    0    0       0     29      0         0      0      0
  Mod       0   13 3057    31    0    0       0      0     16         0      0      0
  PrePr     0    0    1   679   33    0       0      0      0         1      0      0
  Pro       0    0    0     0 1256   32       0      0      0         0      5      0
  PhC       0    0    0     0    0 4597       0      0      0         0      0     19
  Sum   10601 9329 3112   710 1289 4629      23     29     16         1      5     19

Transitions:
     To
From     Records:  Events: Risk time:  Persons:
  None      10689       99     873.08       115
  Min        9334       94     757.05       100
  Mod        3117       60     250.40        63
  PrePr       714       35      56.00        35
  Pro        1293       37     103.99        39
  PhC        4616       19     377.07        36
  Sum       29763      344    2417.60       160
```

In order to get the transitions right we must redefine the up and down transitions by the newly formed split records of follow-up:

```
> sret <- transform( sret, tr.up = as.integer(lex.Cst)<as.integer(lex.Xst) &
+                                  lex.Xst %in% llev,
+                          tr.dn = as.integer(lex.Cst)>as.integer(lex.Xst) &
+                                  lex.Xst %in% llev )
> ftable( addmargins( xtabs( cbind(up=tr.up,dn=tr.dn) ~ lex.Cst + allocation,
+                            data=sret )[1:6,,],
+                     margin=1:2 ),
+         col.vars=3:2 )
                          up                          dn
        allocation Intensive Conventional Sum Intensive Conventional Sum
lex.Cst
None                      37           39  76         0            0   0
Min                       20           34  54         9            2  11
Mod                       11           20  31         8            5  13
PrePr                     13           20  33         1            0   1
Pro                       13           19  32         0            0   0
PhC                        0            0   0         0            0   0
Sum                       94          132 226        18            7  25
```

We note that the number of improvements (`dn`) is quite thin; only 25 transitions recorded, of which a mere 7 in the conventional group.

Then we can fit models for each of the transitions, first we determine the position of the knots for the underlying splines:

```
> with( subset( sret, tr.up ), quantile( tfe+lex.dur, probs=0:3/4 ) )
       0%       25%       50%       75%
0.0428767 2.0244398 4.8444257 7.6056036
```

```
> with( subset( sret, tr.dn ), quantile( tfe+lex.dur, probs=0:3/4 ) )
       0%       25%       50%       75%
0.1170675 1.4905300 5.0865665 9.8096788
> with( subset( sret, lex.Xst %in% dlev), quantile( tfe+lex.dur, probs=0:3/4 ) )
       0%       25%       50%       75%
 1.637235  7.151266 10.193018 15.863107
```

Thus we use the same set of knots for the changes in retinopathy status:

```
> ( r.kn <- c( 0, (
+   with(subset(sret,tr.up), quantile( tfe+lex.dur, probs=1:3/4 )) +
+   with(subset(sret,tr.dn), quantile( tfe+lex.dur, probs=1:3/4 )))/2 ) )
             25%      50%      75%
0.000000 1.757485 4.965496 8.707641
> ( d.kn <- with( subset( sret, lex.Xst %in% dlev ),
+                 quantile( tfe+lex.dur, probs=(1:4-0.5)/4 ) ) )
    12.5%     37.5%     62.5%     87.5%
 4.960986  8.870637 12.206708 16.810404
```

We can now fit models for increase and decrease in retinopathy score as well as mortality to the split dataset, using a smooth underlying hazard. We simplify things by using only linear interactions between times and state (non-proportionality). We make a test for interaction between state and time (removing it from the model), and a test for interaction between state and allocation (adding it to the model):

```
> # Transitions to worse retinopathy
> up.i <- glm( tr.up ~ Ns( tfe, knots=r.kn) + lex.Cst + lex.Cst:tfe + allocation + age + sex,
+              offset = log(lex.dur),
+              family = poisson,
+                data = subset(sret,lex.Cst %in% llev[1:5]) )
> up.0   <- update( up.i, . ~ . - lex.Cst:tfe )
> up.s   <- update( up.i, . ~ . + lex.Cst:allocation )
> anova( up.0, up.i, up.s, test="Chisq" )[-1,-(1:2)]
  Df Deviance Pr(>Chi)
2  4   22.162 0.000186
3  4    1.213 0.875953
> # Transitions to better retinopathy status --- we leave out the
> # interaction for want of data
> dn.0 <- glm( tr.dn ~ Ns( tfe, knots=r.kn) + lex.Cst + allocation + age + sex,
+              offset = log(lex.dur),
+              family = poisson,
+                data = subset(sret,lex.Cst %in% llev[2:6]) )
> dn.s   <- update( dn.0, . ~ . + lex.Cst:allocation )
> anova( dn.0, dn.s, test="Chisq" )[-1,-(1:2)]
  Df Deviance Pr(>Chi)
2  4   1.2079   0.8768
> # Transitions to death
> mt.i <- update( up.i, (lex.Xst %in% levels(lex.Xst)[7:12]) ~ .,
+                       data = sret )
> mt.0   <- update( mt.i, . ~ . - lex.Cst:tfe )
> mt.s   <- update( mt.i, . ~ . + lex.Cst:allocation )
> anova( mt.0, mt.i, mt.s, test="Chisq" )[-1,-(1:2)]
  Df Deviance Pr(>Chi)
2  5   0.2236  0.99884
3  5   9.5370  0.08947
```

For the transitions to worse retinopathy, we see that there is an interaction between state and time, but none for mortality. For the transitions to better retinopathy status we were not able to fit a model with state by time interaction because of the rather thin data, and we saw no indication of interaction between state and allocation (also because of the rather thin data). Thus the relevant model for the deterioration is `up.i`, for improvement `dn.0`, and for mortality `mt.0`.

```
> round( ci.exp(up.i,pval=T), 3 )
                        exp(Est.)  2.5% 97.5%     P
(Intercept)                 0.179 0.053 0.601 0.005
Ns(tfe, knots = r.kn)1      2.965 1.536 5.725 0.001
Ns(tfe, knots = r.kn)2      1.857 0.453 7.607 0.390
Ns(tfe, knots = r.kn)3      1.780 0.843 3.758 0.130
lex.CstMin                  0.726 0.388 1.356 0.315
lex.CstMod                  0.793 0.339 1.854 0.592
lex.CstPrePr                1.759 0.808 3.829 0.155
lex.CstPro                  1.200 0.528 2.724 0.663
allocationConventional      1.311 0.999 1.720 0.051
age                         0.998 0.979 1.018 0.839
sexMale                     1.112 0.819 1.510 0.498
lex.CstNone:tfe             0.809 0.715 0.915 0.001
lex.CstMin:tfe              0.856 0.758 0.967 0.012
lex.CstMod:tfe              0.912 0.794 1.048 0.193
lex.CstPrePr:tfe            1.027 0.910 1.158 0.668
lex.CstPro:tfe              1.000 1.000 1.000   NaN
> round( ci.exp(dn.0,pval=T), 3 )
                        exp(Est.)  2.5% 97.5%     P
(Intercept)                 0.087 0.002 3.461 0.194
Ns(tfe, knots = r.kn)1      0.417 0.123 1.416 0.161
Ns(tfe, knots = r.kn)2      0.023 0.001 0.377 0.008
Ns(tfe, knots = r.kn)3      0.446 0.146 1.364 0.157
lex.CstMod                  3.780 1.667 8.571 0.001
lex.CstPrePr                0.985 0.125 7.779 0.989
lex.CstPro                  0.000 0.000   Inf 0.993
lex.CstPhC                  0.000 0.000   Inf 0.987
allocationConventional      0.294 0.122 0.712 0.007
age                         1.008 0.950 1.070 0.787
sexMale                     1.175 0.392 3.522 0.773
> round( ci.exp(mt.0,pval=T), 3 )
                        exp(Est.)  2.5%      97.5%     P
(Intercept)                 0.000 0.000      0.000 0.000
Ns(tfe, knots = r.kn)1     10.529 1.013    109.485 0.049
Ns(tfe, knots = r.kn)2    108.997 0.091 130456.335 0.195
Ns(tfe, knots = r.kn)3      4.689 1.070     20.555 0.040
lex.CstMin                  0.880 0.500      1.547 0.657
lex.CstMod                  2.203 1.133      4.287 0.020
lex.CstPrePr                0.738 0.098      5.540 0.768
lex.CstPro                  2.099 0.783      5.628 0.141
lex.CstPhC                  1.279 0.678      2.415 0.447
allocationConventional      1.774 1.149      2.740 0.010
age                         1.125 1.085      1.167 0.000
sexMale                     1.497 0.906      2.471 0.115
```

From these we extract the RRs associated with allocation:

```
> HRests <- rbind( ci.exp( up.i, subset="allo", pval=T ),
+                  ci.exp( dn.0, subset="allo", pval=T ),
+                  ci.exp( mt.0, subset="allo", pval=T ) )
> HRests <- cbind( 1/HRests[,c(1,3,2)], HRests )
> rownames( HRests ) <- c("Worse","Better","Death")
> colnames( HRests )[c(1,4)] <- c("Int vs Conv","Conv vs Int")
> round( HRests, 3 )
       Int vs Conv 97.5%  2.5% Conv vs Int  2.5% 97.5%     P
Worse        0.763 0.581 1.001       1.311 0.999 1.720 0.051
Better       3.400 1.405 8.229       0.294 0.122 0.712 0.007
Death        0.564 0.365 0.870       1.774 1.149 2.740 0.010
```

There is a 21% smaller rate of deterioration and a 3 times bigger rate of improvement in retinopathy in the intensive group; the improvement rates very uncertain, though. Controlled for retinopathy state the HR of death is 47% smaller in the intensive group.

5.2.3.3 Progression probabilities

With these models in place we can now set up a transition object; note that for retinopathy progression we use a model with a state by time interaction, and for improvement of retinopathy and mortality we use a proportional hazards model, *i.e.* a model where the intervention has the same effect in all states and in all times:

```
> Tr <- list( "None" = list(  "Min" = up.i,
+                           "D(None)" = mt.0 ),
+              "Min" = list( "None" = dn.0,
+                            "Mod" = up.i,
+                          "D(Min)" = mt.0 ),
+              "Mod" = list(  "Min" = dn.0,
+                           "PrePr" = up.i,
+                          "D(Mod)" = mt.0 ),
+            "PrePr" = list(  "Mod" = dn.0,
+                            "Pro" = up.i,
+                       "D(PrePr)" = mt.0 ),
+             "Pro" = list( "PrePr" = dn.0,
+                            "PhC" = up.i,
+                          "D(Pro)" = mt.0 ),
+            "PhC" = list( "D(PhC)" = mt.0 ) )
> lapply(Tr,names)
$None
[1] "Min"     "D(None)"

$Min
[1] "None"   "Mod"    "D(Min)"

$Mod
[1] "Min"    "PrePr"  "D(Mod)"

$PrePr
[1] "Mod"      "Pro"      "D(PrePr)"

$Pro
[1] "PrePr"  "PhC"    "D(Pro)"

$PhC
[1] "D(PhC)"
```

In order to assess how retinopathy scores have developed in the two groups we simulate how retinopathy score and death is predicted to develop for different types of patient populations; in all case we simulate in two scenarios: Intensive and Conventional.

- A cohort that has the same baseline (joint) distribution of retinopathy scores, age and sex as the Steno 2 patients.
- A cohort where all persons have a retinopathy score of "None", and the same age and sex distribution as the Steno 2 patients.
- Cohorts where all have retinopathy score "None", and the same baseline age (45,50,55,60,65) and sex. This will be simulated using just one cohort with the 10 different combinations represented.

This is basically accomplished by setting up different initial cohorts; `ini.e` corresponding to the empirical distribution of score age and sex, `ini.0` with everyone having a score of 0 but the same age-sex-distribution, and finally `ini.as` with fixed values of age and sex and retinopathy score.

In order to get a Lexis object that corresponds to the baseline state we extract the earliest record for each person; the calendar time will then be the date of baseline, and age the age at baseline:

```
> ini.e <- subset( ret, per==ave(per,lex.id,FUN=min),
+                       select = c("lex.id","lex.Cst",timeScales(ret),"allocation","sex") )
> ini.e <- rbind( transform( ini.e, allocation=factor(levels(allocation)[1]) ),
+                 transform( ini.e, allocation=factor(levels(allocation)[2]) ) )
> str( ini.e )
Classes 'Lexis' and 'data.frame':        320 obs. of  7 variables:
 $ lex.id    : num  1 2 3 4 5 6 7 8 9 10 ...
 $ lex.Cst   : Factor w/ 12 levels "None","Min","Mod",..: 1 2 2 1 1 1 1 2 2 1 ...
 $ per       : num  1993 1993 1993 1993 1993 ...
 $ age       : num  61.1 46.5 50 48.6 57.2 ...
 $ tfe       : num  0 0 0 0 0 0 0 0 0 0 ...
 $ allocation: Factor w/ 2 levels "Intensive","Conventional": 1 1 1 1 1 1 1 1 1 1 ...
 $ sex       : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 1 2 2 2 ...
 - attr(*, "breaks")=List of 3
  ..$ per: NULL
  ..$ age: NULL
  ..$ tfe: NULL
 - attr(*, "time.scales")= chr  "per" "age" "tfe"
 - attr(*, "time.since")= chr  "" "" ""
> with( ini.e, table( allocation, lex.Cst ) )
              lex.Cst
allocation     None Min Mod PrePr Pro PhC D(None) D(Min) D(Mod) D(PrePr) D(Pro) D(PhC)
  Intensive     108  27  10     5   6   4       0      0      0        0      0      0
  Conventional  108  27  10     5   6   4       0      0      0        0      0      0
> ini.0 <- transform( ini.e, lex.Cst=factor( "None", levels=levels(lex.Cst) ) )
> str( ini.0 )
Classes 'Lexis' and 'data.frame':        320 obs. of  7 variables:
 $ lex.id    : num  1 2 3 4 5 6 7 8 9 10 ...
 $ lex.Cst   : Factor w/ 12 levels "None","Min","Mod",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ per       : num  1993 1993 1993 1993 1993 ...
 $ age       : num  61.1 46.5 50 48.6 57.2 ...
 $ tfe       : num  0 0 0 0 0 0 0 0 0 0 ...
 $ allocation: Factor w/ 2 levels "Intensive","Conventional": 1 1 1 1 1 1 1 1 1 1 ...
 $ sex       : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 1 2 2 2 ...
 - attr(*, "breaks")=List of 3
  ..$ per: NULL
  ..$ age: NULL
  ..$ tfe: NULL
 - attr(*, "time.scales")= chr  "per" "age" "tfe"
 - attr(*, "time.since")= chr  "" "" ""
> with( ini.0, table( allocation, lex.Cst ) )
              lex.Cst
allocation     None Min Mod PrePr Pro PhC D(None) D(Min) D(Mod) D(PrePr) D(Pro) D(PhC)
  Intensive     160   0   0     0   0   0       0      0      0        0      0      0
  Conventional  160   0   0     0   0   0       0      0      0        0      0      0
```

With the two initiation cohorts we can now simulate the predicted outcomes according to the model for the transitions:

```
> sim.e  <- simLexis( Tr=Tr, init=ini.e, time.pts=seq(0,22,0.2), N=100 )
> sim.0  <- simLexis( Tr=Tr, init=ini.0, time.pts=seq(0,22,0.2), N=100 )
> save( sim.e, sim.0, file="../data/simRet.Rda" )


> load( file="../data/simRet.Rda" )
> summary( sim.e )
```

```
Transitions:
     To
From    None   Min   Mod PrePr  Pro  PhC D(None) D(Min) D(Mod) D(PrePr) D(Pro) D(PhC)
  None  3670 15294     0     0    0    0    5123      0      0         0      0      0
  Min   2487  4495 11013     0    0    0       0   5357      0         0      0      0
  Mod      0  2658   905  5856    0    0       0      0   3788         0      0      0
  PrePr    0     0   194    86 6339    0       0      0      0       237      0      0
  Pro      0     0     0     0  295 6085       0      0      0         0   1159      0
  PhC      0     0     0     0    0 3019       0      0      0         0      0   3866
  Sum   6157 22447 12112  5942 6634 9104    5123   5357   3788       237   1159   3866

Transitions:
     To
From     Records:  Events: Risk time:  Persons:
  None      24087    20417  174119.59     22618
  Min       23352    18857  165896.96     20629
  Mod       13207    12302   52634.28     12213
  PrePr      6856     6770   10922.26      6772
  Pro        7539     7244   19377.49      7539
  PhC        6885     3866   74531.31      6885
  Sum       81926    69456  497481.89     32000
```

```
> summary( sim.0 )
```

```
Transitions:
     To
From    None   Min   Mod PrePr  Pro  PhC D(None) D(Min) D(Mod) D(PrePr) D(Pro) D(PhC)
  None  5110 21972     0     0    0    0    7041      0      0         0      0      0
  Min   2123  5214 10445     0    0    0       0   5940      0         0      0      0
  Mod      0  1750   896  4376    0    0       0      0   3508         0      0      0
  PrePr    0     0    85    71 4033    0       0      0      0       187      0      0
  Pro      0     0     0     0  278 2931       0      0      0         0    824      0
  PhC      0     0     0     0    0 1478       0      0      0         0      0   1453
  Sum   7233 28936 11426  4447 4311 4409    7041   5940   3508       187    824   1453

Transitions:
     To
From     Records:  Events: Risk time:  Persons:
  None      34123    29013  248461.79     32000
  Min       23722    18508  172455.13     21444
  Mod       10530     9634   42326.30     10033
  PrePr      4376     4305    6105.24      4346
  Pro        4033     3755   10132.37      4033
  PhC        2931     1453   26116.25      2931
  Sum       79715    66668  505597.08     32000
```

5.2.3.4 Prediction from baseline retinopathy distribution

Here we show how a population of patients with an age- sex- and retinopathy-score distribution as the Steno 2 patients would fare through the states were they subjected to intensive, respectively conventional treatment.

```
> prm <- c(1:6,12:7)
> prm
 [1]  1  2  3  4  5  6 12 11 10  9  8  7
```

```
> levels( sim.e )[prm]
```

```
 [1] "None"     "Min"      "Mod"      "PrePr"    "Pro"      "PhC"      "D(PhC)"
 [8] "D(Pro)"   "D(PrePr)" "D(Mod)"   "D(Min)"   "D(None)"
```

```
> rEi <- pState( nState( subset(sim.e,allocation=="Intensive"),
+                        at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=prm )
> rEc <- pState( nState( subset(sim.e,allocation=="Conventional"),
+                        at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=prm )
> round( head( rEi ), 4 )
```

```
      State
when     None    Min    Mod  PrePr    Pro    PhC D(PhC) D(Pro) D(PrePr) D(Mod) D(Min)
  0    0.6750 0.8438 0.9062 0.9375 0.9750 1.0000 1.0000 1.0000   1.0000 1.0000 1.0000
  0.25 0.6541 0.8471 0.9046 0.9349 0.9734 0.9996 0.9997 0.9997   0.9997 0.9998 0.9998
  0.5  0.6352 0.8466 0.9034 0.9329 0.9710 0.9994 0.9994 0.9996   0.9996 0.9997 0.9998
  0.75 0.6178 0.8440 0.9019 0.9304 0.9690 0.9990 0.9991 0.9992   0.9993 0.9994 0.9996
  1    0.6000 0.8420 0.9001 0.9276 0.9664 0.9982 0.9983 0.9986   0.9986 0.9988 0.9991
  1.25 0.5844 0.8378 0.8979 0.9238 0.9633 0.9972 0.9973 0.9976   0.9977 0.9981 0.9984
      State
when   D(None)
  0          1
  0.25       1
  0.5        1
  0.75       1
  1          1
  1.25       1
```

Then plot the estimated distributions over the

```
> retst <- function( Pi, Pc, rv=FALSE, leg=rep(15,12) )
+ {
+ par( mfrow=c(1,2), mar=c(0,3,0,0), oma=c(3,0,2,3-2*rv), las=1 )
+ plot( Pi, col=clr[prm], xlim=c(0,20), xlab="", # yaxt="n",
+           col.lab="transparent" )
+ lines( as.numeric(rownames(Pi)), Pi[,"PhC"], lwd=4 )
+ Pi <- cbind(0,Pi)
+ wr <- cbind( match(paste(leg),rownames(Pi)), 2:13 )
+ wl <- cbind( match(paste(leg),rownames(Pi)), 1:12 )
+ if(is.numeric(leg)) text( leg,
+                           (Pi[wl]+Pi[wr])/2,
+                           colnames(Pi)[-1],
+                           col=rep(c("white","black"),each=6) )
+ axis( side=4, at=0:5/5 )
+ axis( side=4, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4, at=0:100  /100, tcl=-0.2, labels=NA )
+ mtext( "Probability", side=2, line=1, las=0 )
+ mtext( "Intensive", side=3, line=1 )
+
+ plot( Pc, col=clr[prm], xlim=if(rv) c(20,0) else c(0,20),
+       xlab="", yaxt="n", col.lab="transparent" )
+ lines( as.numeric(rownames(Pc)), Pc[,"PhC"], lwd=4 )
+ axis( side=4-2*rv, at=1:5/5, if(rv) labels=NA )
+ axis( side=4-2*rv, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4-2*rv, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4-2*rv, at=0:100  /100, tcl=-0.2, labels=NA )
+ mtext( "Conventional", side=3, line=1 )
+ mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
+ }
> retst(rEi,rEc,rv=FALSE,leg=c(6:8,5,10,12,18,19,rep(18,4)) )
```

We can also plot it reversely for the conventional group to facilitate comparison

```
> retst(rEi,rEc,rv=TRUE,leg=c(6:8,6,9,12,18,19,rep(18,4)) )
```

5.2.3.5 Prediction given absence of retinopathy

Here we show how a population of patients with an age- and sex-distribution as the Steno 2 patients all without retinopathy at baseline, would fare through the states were they subjected to intensive, respectively conventional treatment.

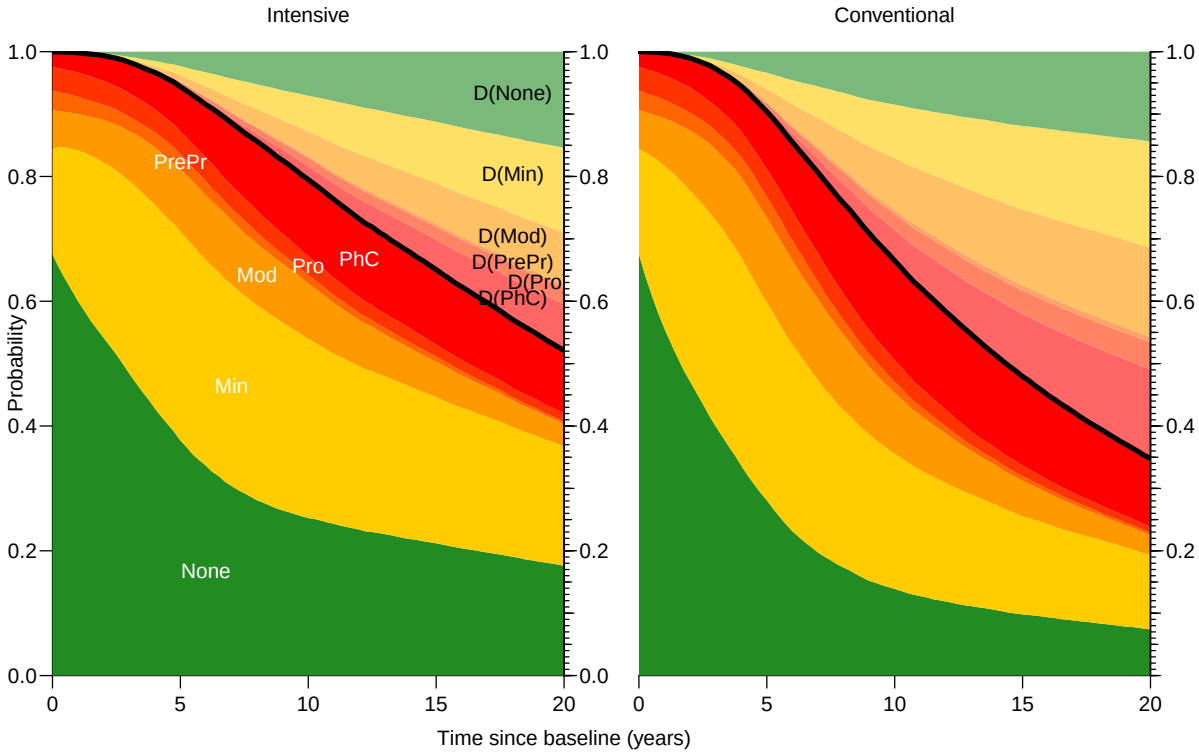

Figure 5.3: *Patients with age- sex- and retinopathy score distribution as in Steno 2 study: Probability of being in each state of retinopathy as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*

```
> rZi <- pState( nState( subset(sim.0,allocation=="Intensive"),
+                       at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=prm )
> rZc <- pState( nState( subset(sim.0,allocation=="Conventional"),
+                       at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=prm )
> round( head( rZi ), 4 )
      State
when     None    Min    Mod  PrePr    Pro    PhC D(PhC) D(Pro) D(PrePr) D(Mod) D(Min)
  0    1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000   1.0000 1.0000 1.0000
  0.25 0.9584 0.9996 0.9999 0.9999 0.9999 0.9999 0.9999 0.9999   0.9999 0.9999 0.9999
  0.5  0.9219 0.9976 0.9996 0.9996 0.9996 0.9996 0.9996 0.9996   0.9996 0.9996 0.9996
  0.75 0.8906 0.9956 0.9991 0.9991 0.9992 0.9992 0.9992 0.9992   0.9992 0.9992 0.9992
  1    0.8602 0.9912 0.9983 0.9985 0.9986 0.9986 0.9986 0.9986   0.9986 0.9986 0.9988
  1.25 0.8348 0.9885 0.9975 0.9978 0.9978 0.9979 0.9979 0.9979   0.9979 0.9979 0.9982
      State
when   D(None)
  0          1
  0.25       1
  0.5        1
  0.75       1
  1          1
  1.25       1
```

Then plot the estimated distributions over the

```
> retst(rZi,rZc,rv=FALSE,leg=c(6:8,6,9,12,18,19,rep(18,4)) )
```

We can also reverse the x-axis for the conventional group to facilitate comparison

```
> retst(rZi,rZc,rv=TRUE,leg=c(6:8,6,9,12,18,19,rep(18,4)) )
```

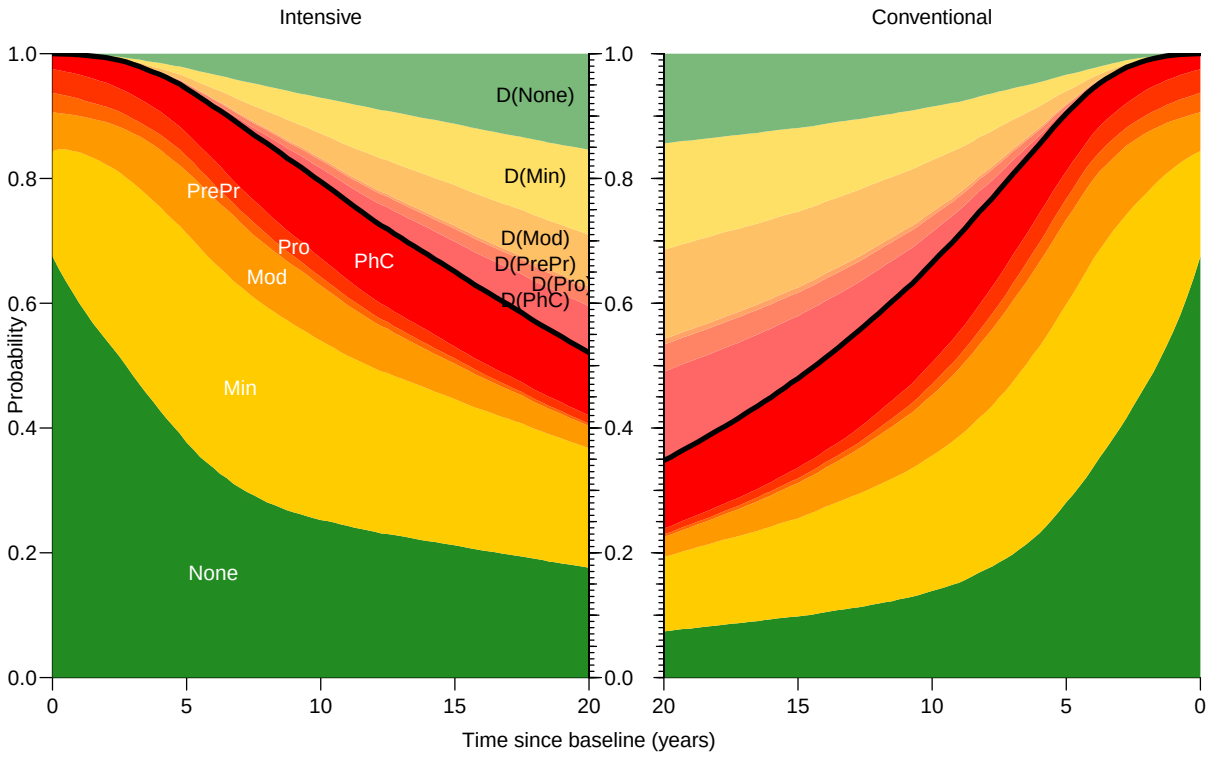

Figure 5.4: *Patients with age- sex- and retinopathy score distribution as in Steno 2 study: Probability of being in each state of retinopathy as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*
This is the same figure as 5.3, except that the x-axis of the right-hand panel is reversed.

5.2.4 Non-reversible retinopathy

In figure 5.2 we see that the the number of transitions to a better state is quite limited, which means that including these very uncertain transitions in the predictions makes them quite uncertain, By this token we also make an analysis of data where we disallow improvement; that is redefining the states to "the worst state seen so far".

We again set the random number generator seed explicitly so that results are reproducible

```
> set.seed( 4667571 )
> ret <- make.dlex( rt, rd, incr=TRUE )
> levels( ret )
 [1] "1"       "2"       "3"       "4"       "5"       "6"       "Dead(1)" "Dead(2)"
 [9] "Dead(3)" "Dead(4)" "Dead(5)" "Dead(6)"

> levels( mic$eds_1 )

[1] "No retinopathy"            "Minimal non-proliferative" "Moderate nonproliferative"
[4] "Pre-proliferative"         "Proliferative"             "Photocoagulation"

> newl <- c("None","Min","Mod","PrePr","Pro","PhC")
> levels( ret$lex.Cst ) <-
+ levels( ret$lex.Xst ) <- c( newl, paste("D(",newl,")",sep="") )
> levels( ret )
 [1] "None"     "Min"      "Mod"      "PrePr"    "Pro"      "PhC"      "D(None)"
 [8] "D(Min)"   "D(Mod)"   "D(PrePr)" "D(Pro)"   "D(PhC)"

> summary( ret )
```

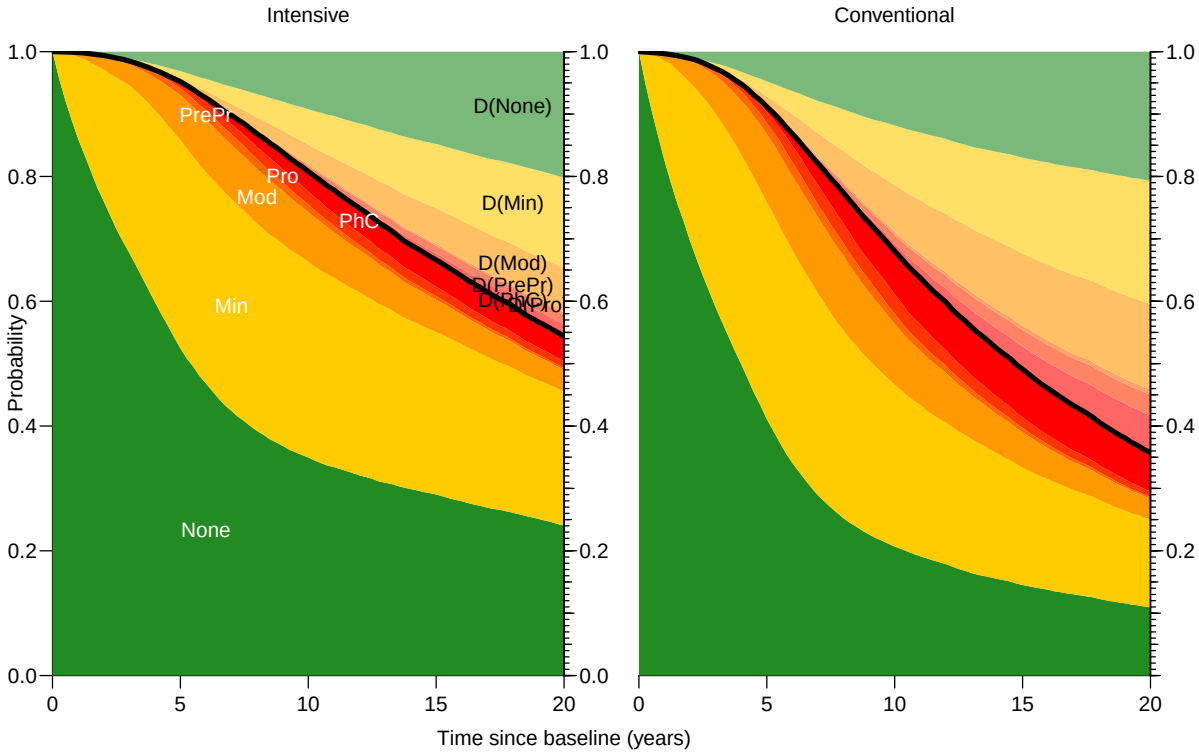

Figure 5.5: *Patients all initiating with retinopathy score 0; age- and sex-distribution as in Steno 2 study population: Probability of being in each state of retinopathy as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*

```
Transitions:
     To
From    None Min Mod PrePr Pro PhC D(None) D(Min) D(Mod) D(PrePr) D(Pro) D(PhC)  Records:
  None   159  70   0     0   0   0      21      0      0         0       0       0       250
  Min      0 156  52     0   0   0       0     27      0         0       0       0       235
  Mod      0   0  78    30   0   0       0      0     20         0       0       0       128
  PrePr    0   0   0     8  33   0       0      0      0         1       0       0        42
  Pro      0   0   0     0  17  32       0      0      0         0       5       0        54
  PhC      0   0   0     0   0  73       0      0      0         0       0      19        92
  Sum    159 226 130    38  50 105      21     27     20         1       5      19       801

Transitions:
     To
From     Events: Risk time:  Persons:
  None        91     806.73       108
  Min         79     699.02        97
  Mod         50     372.22        62
  PrePr       34      72.65        35
  Pro         37      97.61        39
  PhC         19     369.36        36
  Sum        310    2417.60       160
```

With this in place we can now show the transitions through retinopathy states in the two groups:

```
> par( mfrow=c(1,2), oma=c(0,0,0,0) )
> for( lv in levels(ret$allocation) )
+    {
+ boxes( subset( ret, allocation==lv ),
+        boxpos=list(x=rep(c(20,80),each=6)+7-14*(lv=="Intensive"),
```

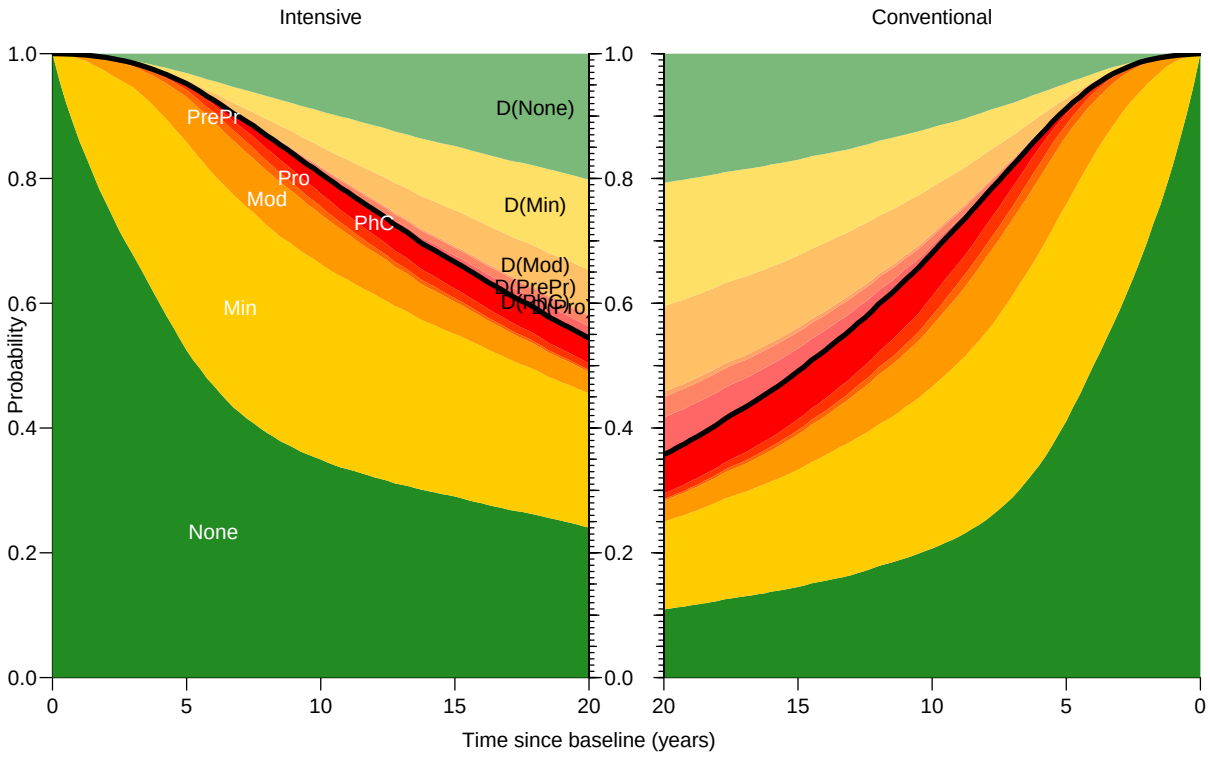

Figure 5.6: *Patients all initiating with retinopathy score 0; age- and sex-distribution as in Steno 2 study population: Probability of being in each state of retinopathy as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*
This is the same figure as 5.5, except that the x-axis of the right-hand panel is reversed.

```
+                    y=rep(seq(90,5,,6),2)),
+         show.BE="noz", scale.R=100,
+         col.bg=clr,
+         col.border=c(clr[1:6],rep("black",6)),
+         col.txt=rep(c("white","black"),each=6) )
+ text(10+5-10*(lv=="Intensive"),98,lv,cex=1.5,font=2,adj=0)
+    }
```

When fitting Cox-models we need two help-variables, namely the indicators of increase, resp. decrease in retinopathy score:

```
> ( llev <- levels(ret)[1:6] )
[1] "None"   "Min"    "Mod"    "PrePr" "Pro"    "PhC"
> ( dlev <- levels(ret)[7:12] )
[1] "D(None)"  "D(Min)"   "D(Mod)"   "D(PrePr)" "D(Pro)"   "D(PhC)"
> ret <- transform( ret, tr.up = as.integer(lex.Cst)<as.integer(lex.Xst) &
+                                lex.Xst %in% llev )
> ftable( xtabs( tr.up ~ lex.Cst + lex.Xst + allocation, data=ret )[1:6,1:6,],
+          col.vars=3:2 )
        allocation Intensive                      Conventional
        lex.Xst         None Min Mod PrePr Pro PhC         None Min Mod PrePr Pro PhC
lex.Cst
None                       0  32   0     0   0   0            0  38   0     0   0   0
Min                        0   0  18     0   0   0            0   0  34     0   0   0
Mod                        0   0   0    10   0   0            0   0   0    20   0   0
PrePr                      0   0   0     0  13   0            0   0   0     0  20   0
Pro                        0   0   0     0   0  13            0   0   0     0   0  19
PhC                        0   0   0     0   0   0            0   0   0     0   0   0
```

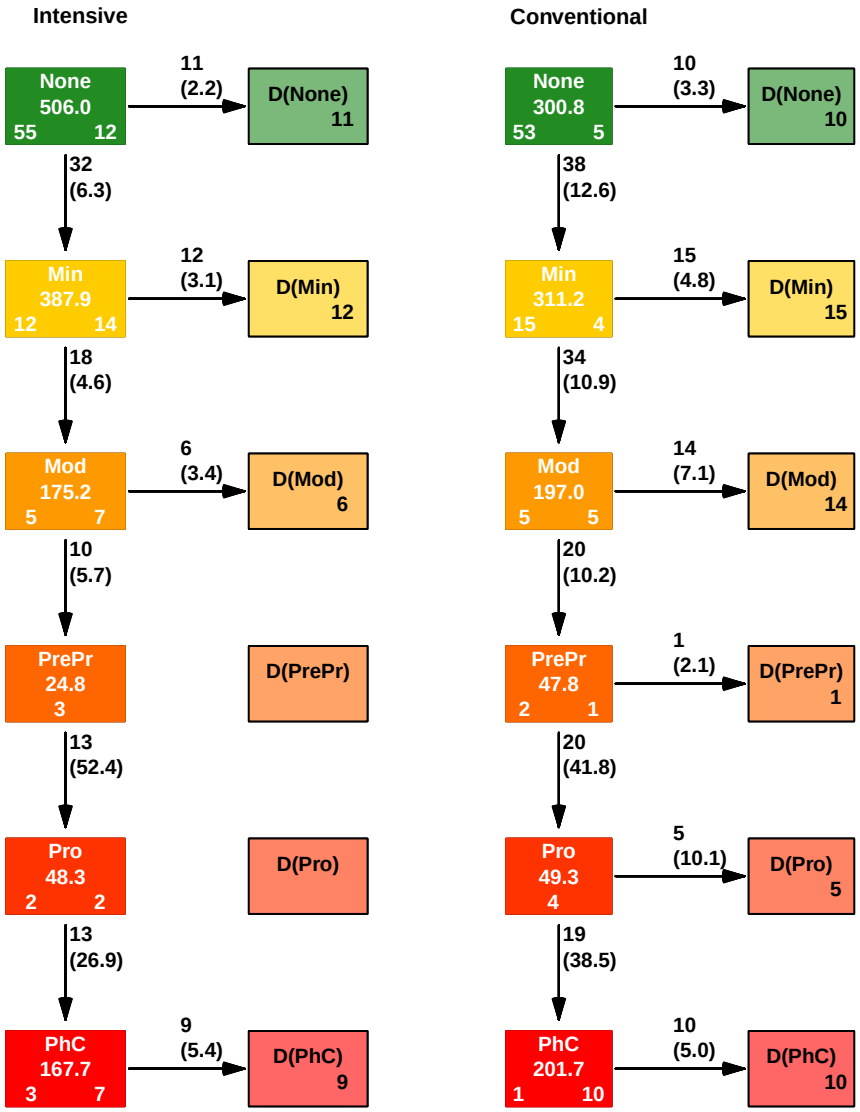

Figure 5.7: *Transitions between states of retinopathy and death allowing only progression of retinopathy. The number in the center of the boxes is the person-years, the numbers at the bottom is the number of patients starting, resp. ending their follow-up in each box, and the numbers on the arrows are the number of transitions and the overall transition rates per 100 PY.*

```
> u.sep <- coxph( Surv( tfe, tfe+lex.dur, tr.up ) ~
+                 strata(lex.Cst) + lex.Cst:allocation,
+                 data = subset(ret,lex.Cst %in% llev[1:5]) )
> u.com <- update( u.sep, . ~ . - lex.Cst:allocation + allocation )
> D.sep <- coxph( Surv( tfe, tfe+lex.dur, lex.Xst %in% dlev ) ~
+                 strata(lex.Cst) + lex.Cst:allocation,
+                 data = ret )
> D.lin <- update( D.sep, . ~ . - lex.Cst:allocation + allocation + allocation:as.integer(lex.Cst) )
> D.com <- update( D.sep, . ~ . - lex.Cst:allocation + allocation )
```

We can check the resulting estimates;

```
> round( ci.exp(u.sep), 3 )
                                  exp(Est.)  2.5% 97.5%
lex.CstNone:allocationIntensive       0.634 0.395 1.019
lex.CstMin:allocationIntensive        0.641 0.360 1.141
lex.CstMod:allocationIntensive        0.691 0.321 1.486
lex.CstPrePr:allocationIntensive      1.098 0.477 2.527
lex.CstPro:allocationIntensive        1.060 0.516 2.179
> round( ci.exp(D.sep), 3 )
                                  exp(Est.)  2.5% 97.5%
lex.CstNone:allocationIntensive       0.517 0.217 1.229
lex.CstMin:allocationIntensive        0.497 0.223 1.103
lex.CstMod:allocationIntensive        0.524 0.199 1.377
lex.CstPrePr:allocationIntensive      0.000 0.000   Inf
lex.CstPro:allocationIntensive        0.000 0.000   Inf
lex.CstPhC:allocationIntensive        1.128 0.457 2.786
```

We use a Wald test to check whether the effect of intervention is the same across categories:

```
> ( CC <- t(matrix( c(1,-1,rep(0,5)), 6, 5 ) ) )
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1   -1    0    0    0    0
[2,]    0    1   -1    0    0    0
[3,]    0    0    1   -1    0    0
[4,]    0    0    0    1   -1    0
[5,]    0    0    0    0    1   -1
> Wald( u.sep, ctr.mat=CC[1:4,1:5] )
    Chisq      d.f.         P
2.4888772 4.0000000 0.6466286
> Wald( D.sep, ctr.mat=CC[1:5,1:6] )
    Chisq      d.f.         P
2.2657049 5.0000000 0.8112905
> round( ci.exp(D.lin,pval=TRUE), 3 )
                                         exp(Est.)  2.5% 97.5%     P
allocationConventional                       2.469 1.097 5.559 0.029
allocationIntensive:as.integer(lex.Cst)      1.117 0.885 1.409 0.351
```

We see that there is no evidence that the HRs differ across the states, thus that the HR of increase between intensive and conventional is constant across retinopathy states, as is the mortality HR. Thus based on the Cox-models we can report the three HRs associated with being in the intensive group:

```
> HR <- rbind( ci.exp(u.com,pval=TRUE),
+              ci.exp(D.com,pval=TRUE) )
> rownames( HR ) <- c("Increasing ret","Death")
> HR <- cbind( 1/HR[,c(1,3,2)], HR )
> colnames(HR)[c(1,4)] <- c("Int vs Conv","Conv vs Int")
> round( HR, 3 )
               Int vs Conv 97.5%  2.5% Conv vs Int  2.5% 97.5%     P
Increasing ret       0.739 0.557 0.980       1.353 1.020 1.795 0.036
Death                0.560 0.363 0.864       1.785 1.158 2.751 0.009
```

Thus we see that there is a 26% smaller HR of increasing retinopathy status in the intensive group (HR 0.74; 0.56–0.99), and a lower mortality when controlling for retinopathy status (HR 0.50; 0.31–0.80), the latter again pretty much confirming the overall analyses of mortality.

5.2.4.1 Parametric modeling

In order to get a more credible model for the retinopathy changes and the mortality, we fit a corresponding set of Poisson models where we also take age and sex into account. To this end we split the follow-up data in 1-month intervals:

```
> timeScales( ret )
[1] "per" "age" "tfe"
> sret <- splitLexis( ret, seq(0,24,1/12), "tfe" )
> summary( sret )
Transitions:
     To
From    None  Min  Mod PrePr  Pro  PhC D(None) D(Min) D(Mod) D(PrePr) D(Pro) D(PhC)
  None  9782   70    0     0    0    0      21      0      0         0      0      0
  Min      0 8533   52     0    0    0       0     27      0         0      0      0
  Mod      0    0 4538    30    0    0       0      0     20         0      0      0
  PrePr    0    0    0   879   33    0       0      0      0         1      0      0
  Pro      0    0    0     0 1183   32       0      0      0         0      5      0
  PhC      0    0    0     0    0 4503       0      0      0         0      0     19
  Sum   9782 8603 4590   909 1216 4535      21     27     20         1      5     19

Transitions:
     To
From     Records:  Events: Risk time:  Persons:
  None        9873       91     806.73       108
  Min         8612       79     699.02        97
  Mod         4588       50     372.22        62
  PrePr        913       34      72.65        35
  Pro         1220       37      97.61        39
  PhC         4522       19     369.36        36
  Sum        29728      310    2417.60       160
```

In order to get the transitions right we must redefine the up transition indicator in the newly formed split records of follow-up:

```
> sret <- transform( sret, tr.up = as.integer(lex.Cst)<as.integer(lex.Xst) &
+                                  lex.Xst %in% llev )
```

Then we can fit models for each of the transitions, first we determine the position of the knots for the underlying splines:

```
> with( subset( sret, tr.up ), quantile( tfe+lex.dur, probs=0:3/4 ) )
        0%        25%        50%        75%
0.04100879 2.05938196 5.12452553 7.70114052
> with( subset( sret, lex.Xst %in% dlev), quantile( tfe+lex.dur, probs=0:3/4 ) )
       0%       25%       50%       75%
 1.637235  7.151266 10.193018 15.863107
```

Thus we use knots for the changes in retinopathy status where we have 0 as the first knot, and for mortality only the quantile knots:

```
> ( r.kn <- c( 0,
+   with(subset(sret,tr.up), quantile( tfe+lex.dur, probs=1:3/4 )) ) )
              25%      50%      75%
0.000000 2.059382 5.124526 7.701141
> ( d.kn <- with( subset( sret, lex.Xst %in% dlev ),
+                 quantile( tfe+lex.dur, probs=(1:4-0.5)/4 ) ) )
    12.5%     37.5%     62.5%     87.5%
 4.960986  8.870637 12.206708 16.810404
```

We can now fit models for increase and decrease in retinopathy score as well as mortality to the split dataset, using a smooth underlying hazard. We simplify things by using only linear interactions between times and state (non-proportionality). We make a test for interaction between states (removing it from the model), and a test for interaction between state and allocation effect (adding it to the model):

```
> # Transitions to worse retinopathy
> up.i <- glm( tr.up ~ Ns( tfe, knots=r.kn) + lex.Cst + lex.Cst:tfe + allocation + age + sex,
+              offset = log(lex.dur),
+              family = poisson,
+                data = subset(sret,lex.Cst %in% llev[1:5]) ) # exclude
>                                                    # risk in the PhC state
> up.0   <- update( up.i, . ~ . - lex.Cst:tfe )
> up.s   <- update( up.i, . ~ . + lex.Cst:allocation )
> anova( up.0, up.i, up.s, test="Chisq" )[-1,-(1:2)]
  Df Deviance  Pr(>Chi)
2  4  23.7004 9.172e-05
3  4   1.9143    0.7515
> # Transitions to death
> mt.i <- update( up.i, (lex.Xst %in% levels(lex.Xst)[7:12]) ~ .,
+                       data = sret )
> mt.0   <- update( mt.i, . ~ . - lex.Cst:tfe )
> mt.s   <- update( mt.i, . ~ . + lex.Cst:allocation )
> anova( mt.0, mt.i, mt.s, test="Chisq" )[-1,-(1:2)]
  Df Deviance Pr(>Chi)
2  5   1.0578  0.95777
3  5   9.5125  0.09029
```

For the transitions to worse retinopathy, we see that there is an interaction between state and time, but none between mortality state and allocation, while neither of these are present for mortality rates.

We then update the HR collector for the retinopaty

```
> up.z   <- update( up.0, . ~ . - age - sex )
> zz <- rbind( ci.exp( up.0, subset="allo", pval=TRUE ),
+              ci.exp( up.z, subset="allo", pval=TRUE ) )
> zz[,1:3] <- 1/zz[,c(1,3,2)]
> load( file="../data/mainCI.Rda" )
> dimnames( mainCI )[[1]][7]
[1] "Retinopthy progression"
> mainCI[7,,] <- zz
> save( mainCI, file="../data/mainCI.Rda" )
```

With these models in place we can now set up a transition object.

```
> Tr <- list( "None" = list(    "Min" = up.i,
+                             "D(None)" = mt.0 ),
+             "Min" = list(    "Mod" = up.i,
+                              "D(Min)" = mt.0 ),
+             "Mod" = list(  "PrePr" = up.i,
+                              "D(Mod)" = mt.0 ),
+          "PrePr" = list(    "Pro" = up.i,
+                           "D(PrePr)" = mt.0 ),
+             "Pro" = list(    "PhC" = up.i,
+                              "D(Pro)" = mt.0 ),
+             "PhC" = list( "D(PhC)" = mt.0 ) )
> lapply(Tr,names)
$None
[1] "Min"     "D(None)"

$Min
[1] "Mod"    "D(Min)"
```

```
$Mod
[1] "PrePr"  "D(Mod)"

$PrePr
[1] "Pro"      "D(PrePr)"

$Pro
[1] "PhC"    "D(Pro)"

$PhC
[1] "D(PhC)"
```

5.2.4.2 Progression probabilities

In order to assess how retinopathy scores have developed in the two groups we simulate how retinopathy score and death is predicted to develop for different types of patient populations; in all case we simulate in two scenarios: Intensive and Conventional.

- A cohort that has the same joint baseline distribution of retinopathy scores, age and sex as the Steno 2 patients.
- A cohort where all persons have a retinopathy score of "None", and the same age and sex distribution as the Steno 2 patients.
- Cohorts where all have retinopathy score "None", and the same baseline age (45,50,55,60,65) and sex. This will be simulated using just one cohort with the 10 different combinations represented.

For the two first scenarios we can use the Lexis objects generated in the previous section, `ini.e` and `ini.0`, respectively.

With the two initiation cohorts we can now simulate the predicted outcomes according to the model for the transitions:

```
> sim.e  <- simLexis( Tr=Tr, init=ini.e , time.pts=seq(0,22,0.1), N=100 )
> sim.0  <- simLexis( Tr=Tr, init=ini.0 , time.pts=seq(0,22,0.1), N=100 )
> save( sim.e, sim.0, file="../data/simRet1.Rda" )

> load( file="../data/simRet1.Rda" )
> summary( sim.e )
Transitions:
     To
From    None   Min   Mod PrePr  Pro  PhC D(None) D(Min) D(Mod) D(PrePr) D(Pro) D(PhC)
  None  3093 13908     0     0    0    0   4599      0      0         0      0      0
  Min      0  4028 10519     0    0    0      0   4761      0         0      0      0
  Mod      0     0  2030  5915    0    0      0      0   4574         0      0      0
  PrePr    0     0     0   205 6491    0      0      0      0       219      0      0
  Pro      0     0     0     0  282 6115      0      0      0         0   1294      0
  PhC      0     0     0     0    0 3037      0      0      0         0      0   3878
  Sum   3093 17936 12549  6120 6773 9152   4599   4761   4574       219   1294   3878

Transitions:
     To
From     Records:  Events: Risk time:  Persons:
  None      21600    18507  160280.17     21600
  Min       19308    15280  151980.20     19308
  Mod       12519    10489   79838.91     12519
  PrePr      6915     6710   15137.05      6915
  Pro        7691     7409   17986.99      7691
  PhC        6915     3878   73866.54      6915
  Sum       74948    62273  499089.86     32000
```

```
> summary( sim.0 )
Transitions:
     To
From    None   Min   Mod PrePr  Pro  PhC D(None) D(Min) D(Mod) D(PrePr) D(Pro) D(PhC)
  None  4534 20525     0     0    0    0    6941      0      0         0      0      0
  Min      0  4769 10091     0    0    0       0   5665      0         0      0      0
  Mod      0     0  1892  4243    0    0       0      0   3956         0      0      0
  PrePr    0     0     0   197 3880    0       0      0      0       166      0      0
  Pro      0     0     0     0  233 2823       0      0      0         0    824      0
  PhC      0     0     0     0    0 1440       0      0      0         0      0   1383
  Sum   4534 25294 11983  4440 4113 4263    6941   5665   3956       166    824   1383

Transitions:
     To
From     Records:  Events: Risk time:  Persons:
  None      32000    27466  237868.87     32000
  Min       20525    15756  163158.34     20525
  Mod       10091     8199   62758.97     10091
  PrePr      4243     4046    9043.21      4243
  Pro        3880     3647    8439.08      3880
  PhC        2823     1383   24603.63      2823
  Sum       73562    60497  505872.10     32000
```

5.2.4.3 Prediction from baseline retinopathy distribution

Here we show how a population of patients with an age- sex- and retinopathy-score distribution as the Steno 2 patients would fare through the states were they subjected to intensive, respectively conventional treatment.

```
> prm <- c(1:6,12:7)
> levels( sim.e )[prm]
 [1] "None"     "Min"      "Mod"      "PrePr"    "Pro"      "PhC"      "D(PhC)"
 [8] "D(Pro)"   "D(PrePr)" "D(Mod)"   "D(Min)"   "D(None)"
> Ei <- pState( nState( subset(sim.e,allocation=="Intensive"),
+                       at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=prm )
> Ec <- pState( nState( subset(sim.e,allocation=="Conventional"),
+                       at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=prm )
> round( head( Ei ), 4 )
      State
when     None    Min    Mod  PrePr    Pro    PhC D(PhC) D(Pro) D(PrePr) D(Mod) D(Min)
  0    0.6750 0.8438 0.9062 0.9375 0.9750 1.0000 1.0000 1.0000   1.0000 1.0000 1.0000
  0.25 0.6444 0.8376 0.9047 0.9354 0.9724 0.9996 0.9996 0.9997   0.9998 0.9998 0.9998
  0.5  0.6198 0.8316 0.9034 0.9336 0.9707 0.9995 0.9995 0.9997   0.9998 0.9998 0.9998
  0.75 0.5986 0.8244 0.9014 0.9312 0.9692 0.9992 0.9992 0.9995   0.9996 0.9996 0.9996
  1    0.5782 0.8184 0.8983 0.9280 0.9664 0.9983 0.9983 0.9986   0.9986 0.9988 0.9989
  1.25 0.5612 0.8121 0.8960 0.9260 0.9641 0.9979 0.9979 0.9982   0.9982 0.9984 0.9987
      State
when   D(None)
  0          1
  0.25       1
  0.5        1
  0.75       1
  1          1
  1.25       1
```

Then we plot the estimated state distribution over the 20 years from baseline

```
> retst(Ei,Ec,rv=FALSE,leg=17)
```

We can also plot it reversely for the conventional group to facilitate comparison

```
> retst(Ei,Ec,rv=TRUE,leg=17)
```

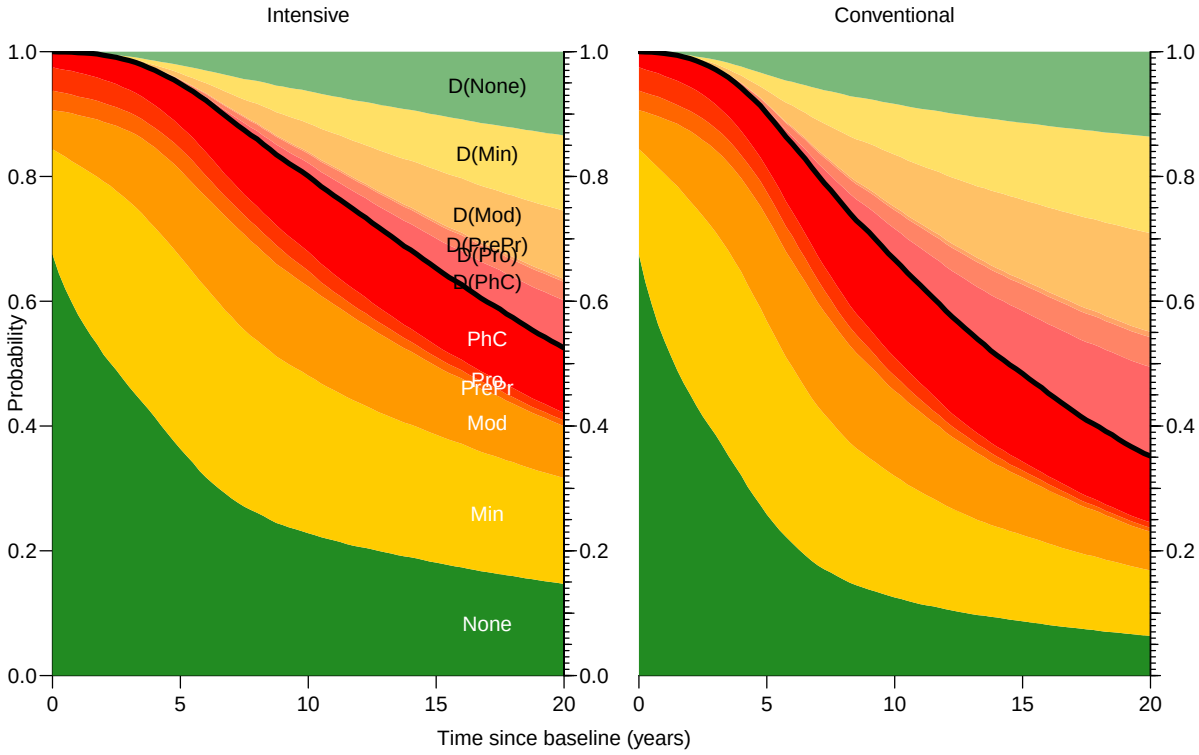

Figure 5.8: *Model with only progression possible: Patients with age- sex- and retinopathy score distribution as in Steno 2 study: Probability of being in each state of retinopathy as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*

5.2.4.4 Prediction given absence of retinopathy

Here we show how a population of patients with an age- and sex-distribution as the Steno 2 patients all without retinopathy at baseline, would fare through the states were they subjected to intensive, respectively conventional treatment.

```
> Zi <- pState( nState( subset(sim.0,allocation=="Intensive"),
+                      at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=prm )
> Zc <- pState( nState( subset(sim.0,allocation=="Conventional"),
+                      at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=prm )
> round( head( Zi ), 4 )
      State
when     None    Min    Mod  PrePr    Pro    PhC D(PhC) D(Pro) D(PrePr) D(Mod) D(Min)
  0    1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000   1.0000 1.0000 1.0000
  0.25 0.9559 0.9984 0.9995 0.9995 0.9995 0.9995 0.9995 0.9995   0.9995 0.9995 0.9995
  0.5  0.9180 0.9969 0.9991 0.9992 0.9992 0.9992 0.9992 0.9992   0.9992 0.9992 0.9992
  0.75 0.8835 0.9945 0.9984 0.9986 0.9986 0.9986 0.9986 0.9986   0.9986 0.9986 0.9986
  1    0.8544 0.9909 0.9977 0.9981 0.9981 0.9981 0.9981 0.9981   0.9981 0.9981 0.9981
  1.25 0.8269 0.9867 0.9968 0.9974 0.9976 0.9976 0.9976 0.9976   0.9976 0.9976 0.9977
      State
when   D(None)
  0          1
  0.25       1
  0.5        1
  0.75       1
  1          1
  1.25       1
```

Then plot the estimated distributions over the

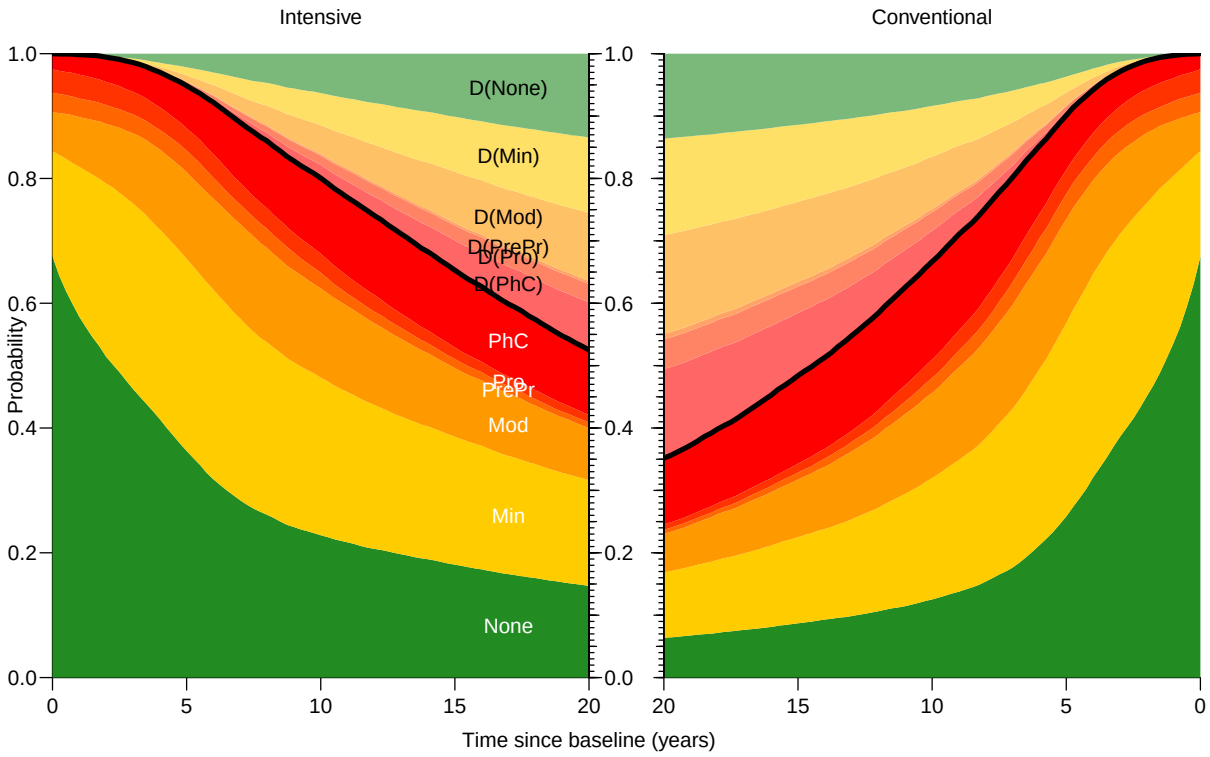

Figure 5.9: *Model with only progression possible: Patients with age- sex- and retinopathy score distribution as in Steno 2 study: Probability of being in each state of retinopathy as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*
This is the same figure as 5.8, except that the x-axis of the right-hand panel is reversed.

```
> retst(Zi,Zc,rv=FALSE,leg=17)
```

We can also reverse the x-axis for the conventional group to facilitate comparison

```
> retst(Zi,Zc,rv=TRUE,leg=17)
```

5.2.4.5 Predictions by age and sex

Finally we make predictions of the development of retinopathy and death for homogeneous groups of patients all starting with retinopathy score 0, and with identical age and sex. The predictions are made for 10 different combinations of age and sex. First we set up a list structure to hold each of the simulated data sets:

```
> lp <- list()
> length( lp ) <- 20
> dim( lp ) <- c(5,2,2)
> dimnames( lp ) <- list( age = seq(45,65,5),
+                         sex = levels(ini.0$sex),
+                         grp = levels(ini.0$allocation)[2:1] )
```

Then we simulate transitions for each combination of age, sex and allocation:

```
> ini <- ini.0[1,]
> system.time(
+ for( ig in dimnames(lp)[[3]] )
+ for( ia in dimnames(lp)[[1]] )
```

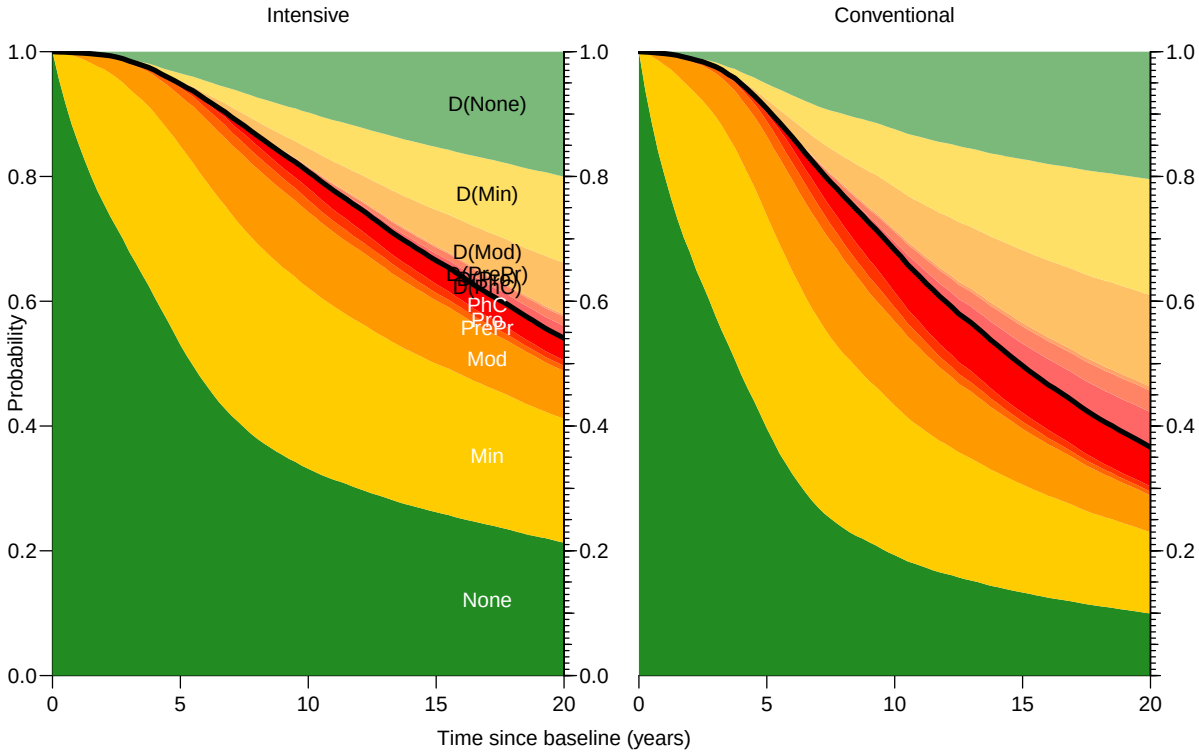

Figure 5.10: *Model with only progression possible, assuming all persons initiating without retinopath and age- and sex-distribution as in Steno 2 study population: Probability of being in each state of retinopathy as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*

```
+ for( is in dimnames(lp)[[2]] )
+    {
+ ini[1,"age"] <- as.numeric(ia)
+ ini[1,"sex"] <- is
+ ini[1,"allocation"] <- ig
+ stp <- simLexis( Tr = Tr,
+                 init = ini,
+                    N = 1000,
+             time.pts = seq(0,21,0.1) )
+ nst <- nState( stp, at=seq(0,20,0.1), from=0, time.scale="tfe" )
+ pst <- pState( nst, perm=prm )
+ lp[[ia,is,ig]] <- pst
+    } )
> save( lp, file="../data/ret1-lp.Rda" )


> load( file="../data/ret1-lp.Rda" )
```

The object `lp` is now a list with 3 dimensions (intervention, age, sex), where each element is a `pState` object, so that we can refer easily to results for any combination of age, sex and allocation.

Using the `lp` object we can now compare the state distribution between persons allocated to intervention and conventional, separately for all ten combinations of age (45, 50, 55, 60, 65) and sex (M, F).

Basically it is the same graph as for the CVD events in the Steno2 population:

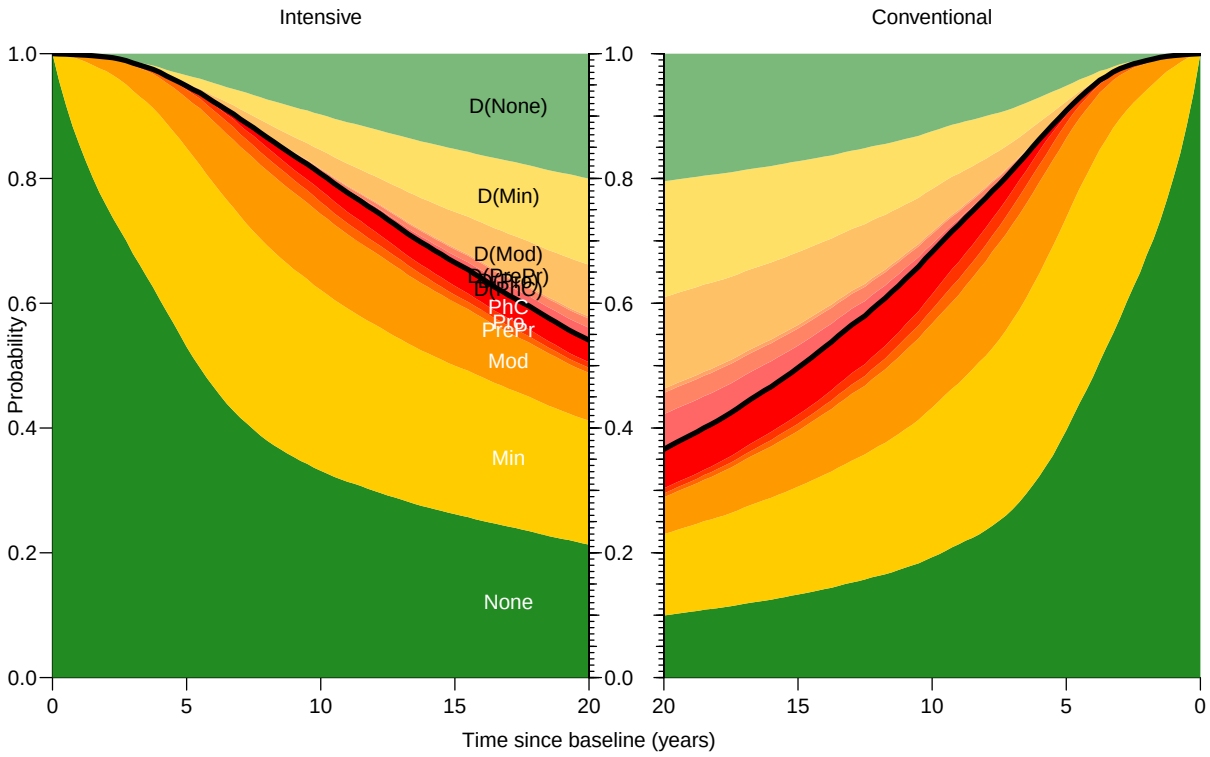

Figure 5.11: *Model with only progression possible, assuming all persons initiating without retinopath and age- and sex-distribution as in Steno 2 study population: Probability of being in each state of retinopathy as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*
This is the same figure as 5.10, except that the x-axis of the right-hand panel is reversed.

```
> pl10 <- function(lp){
+ par( mfrow=c(5,4), mar=c(0,0,0,3), oma=c(4,4,4,0), las=1 )
+ for( ia in dimnames(lp)[["age"]] )
+ for( is in dimnames(lp)[["sex"]][2:1] )
+ for( ig in dimnames(lp)[["grp"]][2:1] )
+    {
+ oax <- ( ig==dimnames(lp)[["grp"]][2] )
+ ptmp <- lp[[ia,is,ig]]
+ plot( ptmp, col=clr[prm], xaxt="n", yaxt="n",
+       xlim=if( oax ) c(0,20) else c(20,0) )
+ if( ia=="45") mtext( ig, side=3, line=1, at=10, outer=FALSE )
+ if( is=="M" & ig=="Intensive" )
+   {
+   axis( side=4-2*oax, at=1:5/5 )
+ # mtext( ia, side=2, line=3, outer=FALSE )
+   }
+ if( is=="F" & ig=="Conventional" ) mtext(ia,side=4,at=0.5,line=0.5,las=1)
+ axis( side=2, labels=NA )
+ if( ia=="65")
+   {
+   axis( side=1 )
+   axis( side=4, at=0, labels="0.0" )
+   if( ia=="65" & is=="M" & oax ) axis( side=2, at=0, labels="0.0" )
+   }
+ lines( as.numeric(rownames(ptmp)), ptmp[,4], col="black", lwd=4 )
+ # text( 1, 0.05, paste( is, ", ", ia, "\n", ig, esp="" ),
+ #       adj=c(1-oax,0), font=2, col="white", cex=1.5 )
```

```
+ axis( side=2+2*oax, at=1:5*20 /100,  if(oax)  labels=NA )
+ axis( side=2+2*oax, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=2+2*oax, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=2+2*oax, at=0:100  /100, tcl=-0.2, labels=NA )
+    }
+ #mtext( "Men"  , side=3, at=0.25, outer=TRUE, line=2 )
+ #mtext( "Women", side=3, at=0.75, outer=TRUE, line=2 )
+ #mtext( "Age at entry", side=3, at=-0.02, outer=TRUE, line=1, adj=0 )
+ mtext( "Time since entry (years)", side=1, outer=TRUE, line=2.5 )
+ mtext( "Probability"             , side=2, outer=TRUE, line=2.5, las=0 )
+ mtext( "Men"  , side=3, at=0.25, outer=TRUE, line=2.5, las=0 )
+ mtext( "Women", side=3, at=0.75, outer=TRUE, line=2.5, las=0 )
+ mtext( "Age"  , side=3, at=1.00, outer=TRUE, line=2.5, las=0, adj=1 )
+ }
> pl10(lp)
```

Finally we save the estimated state distributions:

```
> save( rEi, rEc, rZi, rZc, Ei, Ec, Zi, Zc, file="../data/pRet.Rda" )
```

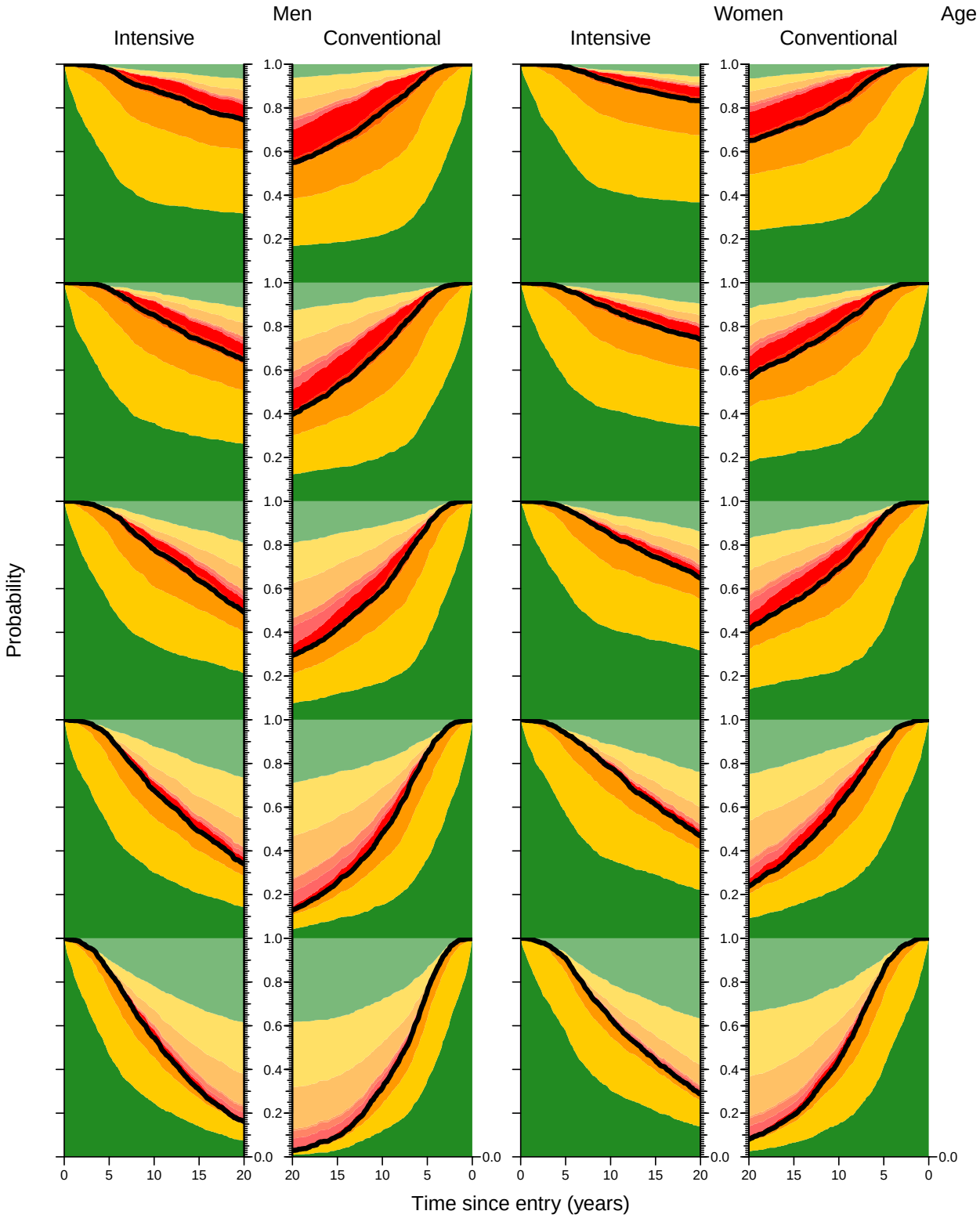

Figure 5.12: *Probabilities of being in different states of retinopathy by age, sex and intervention group. State coloring is the same as in figure* **??**.

5.3 Neuropathy

5.3.1 Data

First we provide an overview of the actually occurring neuropathy patterns in the data, for both types of neuropathy:

```
> library( Epi )
> library( survival )
> clear()
> load( file="../data/mic.Rda" )
> with( mic, ftable(visit,autoprog,periprog,useNA="ifany",col.vars=c(1,3)) )
         visit      1           2           3           4           5           6
         periprog   0   1  NA   0   1  NA   0   1  NA   0   1  NA   0   1  NA   0   1  NA
autoprog
0                   0   0   0   0   0 121  95  20   0  41  30   0  24  16   0  18  12   1
1                   0   0   0   0   0  20  24   8   0  33  19   0  21  22   2  10  12   1
NA                  0   0 160   0   0  15   0   2   0   2   1   2   4   1   3   0   0   8
> with( mic, ftable(addmargins(table(visit,allocation,autoprog,periprog,useNA="ifany"),
+                              margin=2:4),row.vars=c(1,4)) )
               allocation Intensive                Conventional                Sum
               autoprog           0   1  NA Sum            0   1  NA Sum   0   1  NA Sum
visit periprog
1     0                            0   0   0   0            0   0   0   0   0   0   0   0
      1                            0   0   0   0            0   0   0   0   0   0   0   0
      NA                           0   0  80  80            0   0  80  80   0   0 160 160
      Sum                          0   0  80  80            0   0  80  80   0   0 160 160
2     0                            0   0   0   0            0   0   0   0   0   0   0   0
      1                            0   0   0   0            0   0   0   0   0   0   0   0
      NA                          61  10   7  78           60  10   8  78 121  20  15 156
      Sum                         61  10   7  78           60  10   8  78 121  20  15 156
3     0                           53   7   0  60           42  17   0  59  95  24   0 119
      1                            9   3   1  13           11   5   1  17  20   8   2  30
      NA                           0   0   0   0            0   0   0   0   0   0   0   0
      Sum                         62  10   1  73           53  22   1  76 115  32   2 149
4     0                           25  14   0  39           16  19   2  37  41  33   2  76
      1                           19   7   1  27           11  12   0  23  30  19   1  50
      NA                           0   0   1   1            0   0   1   1   0   0   2   2
      Sum                         44  21   2  67           27  31   3  61  71  52   5 128
5     0                           17  15   2  34            7   6   2  15  24  21   4  49
      1                           10   8   1  19            6  14   0  20  16  22   1  39
      NA                           0   1   1   2            0   1   2   3   0   2   3   5
      Sum                         27  24   4  55           13  21   4  38  40  45   8  93
6     0                           15   6   0  21            3   4   0   7  18  10   0  28
      1                            8   7   0  15            4   5   0   9  12  12   0  24
      NA                           0   0   4   4            1   1   4   6   1   1   8  10
      Sum                         23  13   4  40            8  10   4  22  31  23   8  62
```

So there are no asessment of peripheral neuropathy progression until at the third visit. We provide an overview of the patterns of neuropathy progressions over time; visits with no assessment are shown with a "-", no progression with 0 and progression with 1.

```
> pt <- with(mic,tapply( periprog, list(ID,visit), mean ))
> pp <- gsub("NA","-",apply( pt, 1, paste, collapse=" "))
> cbind( sort( addmargins( table( pp ) ) ) )
           [,1]
- - - - 0 -    1
- - 0 0 - 1    1
- - 0 1 0 1    1
- - 0 1 - 1    1
- - - 1 - -    1
- - - 1 - 0    1
- - 1 1 - 1    1
```

```
- - 0 - 0 -    2
- - 0 0 1 0    2
- - 0 1 1 0    2
- - 1 1 1 0    2
- - 0 1 1 1    3
- - 0 0 1 1    4
- - 0 1 0 0    4
- - 0 1 - -    5
- - 1 - - -    5
- - 0 0 0 1    6
- - 0 0 1 -    6
- - 1 1 1 -    6
- - 0 1 1 -    7
- - 1 1 1 1    7
- - - - - -    8
- - 1 1 - -    9
- - 0 0 0 0   17
- - 0 - - -   18
- - 0 0 0 -   18
- - 0 0 - -   22
Sum          160
> at <- with(mic,tapply( autoprog, list(ID,visit), mean ))
> aa <- gsub("NA","-",apply( at, 1, paste, collapse=" "))
> cbind( sort( addmargins( table( aa ) ) ) )
            [,1]
- - - 0 - -    1
- - 0 - - -    1
- 0 0 0 - 1    1
- - 0 0 1 -    1
- 0 0 - 1 -    1
- 0 0 1 - 0    1
- 0 0 1 - 1    1
- 0 0 1 1 0    1
- 0 - 1 - 0    1
- 0 1 - 0 -    1
- 0 1 0 1 0    1
- 0 1 1 0 -    1
- 0 1 1 1 0    1
- - 1 - - -    1
- 1 0 0 0 -    1
- 1 0 0 1 -    1
- - 1 0 1 -    1
- 1 0 1 0 0    1
- 1 0 1 1 0    1
- 1 0 1 1 1    1
- - 1 1 - -    1
- 1 1 0 - -    1
- 1 1 0 0 -    1
- 1 1 - - 1    1
- 1 1 1 0 -    1
- 1 1 1 0 0    1
- 1 1 1 1 1    1
- - 0 0 0 0    2
- 0 0 0 0 1    2
- 0 0 1 0 -    2
- 0 0 1 0 0    2
- 0 0 1 0 1    2
- - 0 1 - -    2
- 0 1 1 1 -    2
- 1 - - - -    2
- 1 0 - - -    2
- 1 1 1 1 -    2
- 1 1 1 - -    3
- 0 - - - -    4
- - 0 0 - -    4
- 0 0 0 1 0    4
```

```
- 0 1 - - -    4
- 0 1 0 - -    4
- 0 1 1 - -    4
- - - - - -    5
- 0 0 1 1 1    5
- 0 0 0 1 -    6
- 0 0 1 1 -    7
- 0 0 0 - -    8
- 0 0 0 0 -    8
- 0 0 1 - -    8
- 0 0 0 1 1    9
- 0 0 - - -   15
- 0 0 0 0 0   15
Sum          160
```

As for the other microvascular outcomes we generate states and dates. Since there are some persons who has no assessments of neuropathy at all we cannot include these in the calculations, so they are excluded *a priori*:

```
> vd <- with( mic, tapply(    doVis, list(ID,visit), mean ) )
> at <- with( mic, tapply( autoprog, list(ID,visit), mean ) )
> pt <- with( mic, tapply( periprog, list(ID,visit), mean ) )
> cbind( at, pt, vd )[1:10,]
    1  2  3  4  5  6  1  2  3  4  5  6        1        2        3        4        5
1  NA  0  0  0  0  0 NA NA  0  0  0  0 1993.357 1995.158 1997.135 2001.097 2006.712
2  NA  1  1  1  1  1 NA NA  0  1  0  1 1993.357 1995.377 1997.781 2001.814 2006.868
3  NA  0  0  1 NA NA NA NA  0  1 NA NA 1993.362 1995.180 1997.151 2001.102       NA
4  NA NA  0  0 NA NA NA NA  1  1 NA NA 1993.362 1995.120 1997.110 2001.066       NA
5  NA  1 NA NA NA NA NA NA  1 NA NA NA 1993.162 1995.002 1997.707       NA       NA
6  NA  1  1  1  1 NA NA NA  0  0  1 NA 1993.201 1995.008 1997.036 2001.025 2006.736
7  NA  0  1  1 NA NA NA NA  0  1 NA NA 1993.203 1995.043 1997.077 2001.058       NA
8  NA  0  1  1  1 NA NA NA  0  0  0 NA 1993.209 1995.829 1997.729 2001.655 2006.903
9  NA  0 NA NA NA NA NA NA NA NA NA NA 1993.209 1995.021       NA       NA       NA
10 NA  0  1  1  1 NA NA NA  0  0  0 NA 1993.075 1995.024 1997.020 2001.023 2006.682
          6
1  2014.802
2  2014.589
3         NA
4         NA
5         NA
6  2015.029
7         NA
8         NA
9         NA
10        NA

> table( wh.a <- apply( at, 1, function(x) any(!is.na(x)) ) )

FALSE  TRUE
    5   155

> table( wh.p <- apply( pt, 1, function(x) any(!is.na(x)) ) )

FALSE  TRUE
    8   152
```

Since we are recording *progression* of neuropathy since baseline, we set the state to 0 at baseline:

```
> at[,1] <- 0
> pt[,1] <- 0
```

5.3.2 Autonomous and peripheral neuropathy

We then set up a proper Lexis objects using `make.dlex` separately for autonomous and peripheral neuropathy. These objects are based on simulated transition times that are uniformly distributed on the intervals in which status changes has taken place, and an assumption of the last registered state being carried over to either death of last recorded visit. Entry is taken as the first visit with a neuropathy measurement. Since the measurements are first made at either 2nd visit (auto) or 3rd visit (peri) follow-up starts at the 2nd visit for autonomous neuropathy, when including peripheral neuropathy we are essentially assuming absence of peripheral neuropathy at 2nd visit.

Since we are simulating an exact transition time we need a random generator seed in order to make the dataset reproducible:

```
> source("usefuns.R")
> set.seed(7471568)
> LxA <- make.dlex( at[wh.a,], vd[wh.a,], incr=TRUE )
> LxP <- make.dlex( pt[wh.p,], vd[wh.p,], incr=TRUE )
> levels(LxA$lex.Cst) <- levels(LxA$lex.Xst) <- c("None","Auto","D(None)","D(Auto)")
> levels(LxP$lex.Cst) <- levels(LxP$lex.Xst) <- c("None","Peri","D(None)","D(Peri)")
> summary( LxA )
Transitions:
     To
From   None Auto D(None) D(Auto)  Records:  Events: Risk time:  Persons:
  None  338   97      37       0       472      134    1319.06       155
  Auto    0  229       0      52       281       52    1058.67        97
  Sum   338  326      37      52       753      186    2377.73       155
> summary( LxP )
Transitions:
     To
From   None Peri D(None) D(Peri)  Records:  Events: Risk time:  Persons:
  None  264   74      53       0       391      127    1536.82       152
  Peri    0  168       0      33       201       33     833.78        74
  Sum   264  242      53      33       592      160    2370.59       152
```

5.3.3 Joint occurrence of autonomous and peripheral neuropathy

In order to set up transitions through states Auto and Peri in combination, we start out with the `LxA` object (the one with the longest risk time), and then cut this at the dates of entry to the Peri state.

```
> lxA <- Relevel( LxA, list("W"=1,"A"=2,"D"=3:4), print=F )
> lxP <- Relevel( LxP, list("W"=1,"P"=2,"D"=3:4), print=F )
> summary( lxA )
Transitions:
     To
From    W   A  D  Records:  Events: Risk time:  Persons:
  W   338  97 37       472      134    1319.06       155
  A     0 229 52       281       52    1058.67        97
  Sum 338 326 89       753      186    2377.73       155
> summary( lxP )
Transitions:
     To
From    W   P  D  Records:  Events: Risk time:  Persons:
  W   264  74 53       391      127    1536.82       152
  P     0 168 33       201       33     833.78        74
  Sum 264 242 86       592      160    2370.59       152
> ( noP <- setdiff( lxA$lex.id, lxP$lex.id ) )
```

```
[1]   9  66  76 134
> ( noA <- setdiff( lxP$lex.id, lxA$lex.id ) )
[1] 60
> length( bothAP <- intersect( lxP$lex.id, lxA$lex.id ) )
[1] 151
```

We note that there are 4 persons with autonomous and no peripheral assessment and one who is vice versa, and 151 patients with at least one assessment of both types of neuropathy. The joint analysis will be based on these persons.

```
> lxA <- subset( lxA, lex.id %in% bothAP )
> lxP <- subset( lxP, lex.id %in% bothAP )
```

From the object with peripheral neuropathy measurements we derive the dates of onset which we shall use to cut the follow-up for autonomous neuropathy. From the peripheral Lexis object we derive the first date of "P":

```
> cutP <- subset( lxP, lex.Cst=="P", select=c("lex.id","per") )
> cutP <- cutP[match(unique(cutP$lex.id),cutP$lex.id),]
> names( cutP )[2] <- "cut"
> head( cutP )
   lex.id      cut
7       2 1999.180
12      3 1998.809
14      4 1996.584
20      5 1993.845
25      6 2003.266
29      7 1998.171
```

Then we cut the records originating in the "W" state, meaning that the new state must be "P":

```
> cutP$new.state = "P"
> lxWP <- cutLexis( subset(lxA,lex.Cst=="W"),
+                   cut=cutP, timescale="per", precursor.states=c("W") )
> summary( subset(lxA,lex.Cst=="W") )
Transitions:
     To
From   W  A  D  Records:  Events: Risk time:  Persons:
   W 335 96 34       465      130    1308.22       151
> summary( lxWP )
Transitions:
     To
From    W   P  A  D  Records:  Events: Risk time:  Persons:
  W   252  46 72 25       395      143     935.70       151
  P     0  83 24  9       116       33     372.51        46
  Sum 252 129 96 34       511      176    1308.22       151
```

However in this object, the transitions labeled from "P" to "A", are really from "P" to "A+P", so we fix this:

```
> lxWP <- transform( lxWP, lex.Xst = ifelse( lex.Cst=="P" & lex.Xst=="A",
+                                            "A+P", as.character(lex.Xst) ) )
> summary( lxWP )
Transitions:
     To
From   A A+P  D   P   W  Records:  Events: Risk time:  Persons:
  W   72   0 25  46 252       395      143     935.70       151
  P    0  24  9  83   0       116       33     372.51        46
  Sum 72  24 34 129 252       511      176    1308.22       151
```

Then we cut the records originating in the "A" state, meaning that the new state must be "A+P"; note that "A" is now taken as a precursor state too:

```
> cutP$new.state = "A+P"
> lxAP <- cutLexis( subset(lxA,lex.Cst=="A"),
+                   cut=cutP, timescale="per", precursor.states=c("W","A") )
> summary( lxAP )
Transitions:
     To
From  W   A A+P  D  Records:  Events: Risk time:  Persons:
  A   0 133  27 28       188       55     598.33        72
  A+P 0   0  95 23       118       23     459.54        51
  Sum 0 133 122 51       306       78    1057.87        96
```

Finally we stack the two and check that it really is an expansion of the `lxA`:

```
> LxAP <- Relevel( rbind( lxAP, lxWP ) )
> levels( LxAP )
[1] "W"   "A"   "A+P" "P"   "D"
```

Ultimately we would also like to see how many died from which state, so we split the Death state:

```
> LxAP <- transform( LxAP, lex.Xst = ifelse( lex.Xst=="D",
+                                            paste("D(",as.character(lex.Cst),")",sep=""),
+                                            as.character(lex.Xst) ) )
> levels( Relevel( LxAP ) )
[1] "W"      "A"      "A+P"    "P"      "D(A)"   "D(A+P)" "D(P)"   "D(W)"
> LxAP <- Relevel( LxAP, c(1,2,4,3,8,5,7,6) )
> levels( LxAP )
[1] "W"      "A"      "P"      "A+P"    "D(W)"   "D(A)"   "D(P)"   "D(A+P)"
> levels( LxAP$lex.Cst) <-
+ levels( LxAP$lex.Xst) <- gsub("W","Well",
+                           gsub("A","Auto",
+                           gsub("P","Peri",levels(LxAP))))
> summary( lxA )
Transitions:
     To
From    W   A  D  Records:  Events: Risk time:  Persons:
  W   335  96 34       465      130    1308.22       151
  A     0 228 51       279       51    1057.87        96
  Sum 335 324 85       744      181    2366.09       151
> summary( LxAP )
Transitions:
     To
From        Well Auto Peri Auto+Peri D(Well) D(Auto) D(Peri) D(Auto+Peri)  Records:
  Well       252   72   46         0      25       0       0            0       395
  Auto         0  133    0        27       0      28       0            0       188
  Peri         0    0   83        24       0       0       9            0       116
  Auto+Peri    0    0    0        95       0       0       0           23       118
  Sum        252  205  129       146      25      28       9           23       817

Transitions:
     To
From         Events: Risk time:  Persons:
  Well           143     935.70       151
  Auto            55     598.33        72
  Peri            33     372.51        46
  Auto+Peri       23     459.54        51
  Sum            254    2366.09       151
```

5.3.4 Collected data for neuropathy

We now have the Lexis objects for separate analysis of autonomous and peripheral neuropathy. For the sake of completeness we merge in the date of diabetes diagnosis and date of baseline from the baseline dataset:

```
> load( "../data/st2.Rda" )
> stb <- st2[,c("ID","doDM","doBase")]
> names( stb )[1] <- "lex.id"
> str( stb )
'data.frame':        160 obs. of  3 variables:
 $ lex.id: num  1 2 3 4 5 6 7 8 9 10 ...
 $ doDM  : num  1991 1982 1983 1977 1986 ...
 $ doBase: num  1993 1993 1993 1993 1993 ...
> LxA  <- merge( LxA , stb, all.x=TRUE )
> LxP  <- merge( LxP , stb, all.x=TRUE )
> LxAP <- merge( LxAP, stb, all.x=TRUE )
```

Then we save the three Lexis objects:

```
> save( LxA, LxP, LxAP, file="../data/LxNeu.Rda" )
```

```
> clr <- c( "forestgreen", heat.colors(5)[c(2,3,1)] )
> clr <- c(clr,rgb(t(col2rgb(clr)*0.6+255*0.4),max=255))
> summary( LxAP )
Transitions:
     To
From         Well Auto Peri Auto+Peri D(Well) D(Auto) D(Peri) D(Auto+Peri)  Records:
  Well        252   72   46         0      25       0       0             0       395
  Auto          0  133    0        27       0      28       0             0       188
  Peri          0    0   83        24       0       0       9             0       116
  Auto+Peri     0    0    0        95       0       0       0            23       118
  Sum         252  205  129       146      25      28       9            23       817

Transitions:
     To
From          Events: Risk time:  Persons:
  Well            143     935.70       151
  Auto             55     598.33        72
  Peri             33     372.51        46
  Auto+Peri        23     459.54        51
  Sum             254    2366.09       151
> par( mfcol=c(1,2), oma=c(0,0,0,0) )
> for( lv in levels(LxAP$allocation) )
+    {
+ boxes( subset( LxAP, allocation==lv ),
+        boxpos=list(x=c(20,30,30,80,50,80,80,80)+7-14*(lv=="Intensive"),
+                    y=c(65,90,10,60,48,90,10,30)),
+        show.BE="noz", scale.R=100,
+       # pos.arr=c(0.45,0.3,0.3),
+        col.bg=clr,
+        col.border=c(clr[1:4],rep("black",4)),
+        col.txt=rep(c("white","black"),each=4) )
+ text(10+5-10*(lv=="Intensive"),98,lv,cex=1.5,font=2,adj=0)
+    }
```

```
> clr <- c( "forestgreen", "orange" )
> clr <- c(clr,rgb(t(col2rgb(clr)*0.6+255*0.4),max=255))
> i <- 0
> par( mfcol=c(2,2), oma=c(0,0,0,0) )
> for( lv in levels(LxA$allocation) )
+ for( Lx in list(LxA,LxP) )
```

```
+    {
+ i <- i + 1
+ boxes( subset( Lx, allocation==lv ),
+        boxpos=list(x=rep(c(20,80),each=2)+7-14*(lv=="Intensive"),
+                    y=rep(c(75,15),2)),
+        show.BE="noz", scale.R=100,
+        pos.arr=c(0.45,0.3,0.3), col.bg=clr,
+        col.border=c(clr[1:2],rep("black",2)),
+        col.txt=rep(c("white","black"),each=2) )
+ if( i %in% c(1,3) ) text(10+5-10*(lv=="Intensive"),98,lv,cex=1.5/0.83,font=2,adj=0)
+    }
```

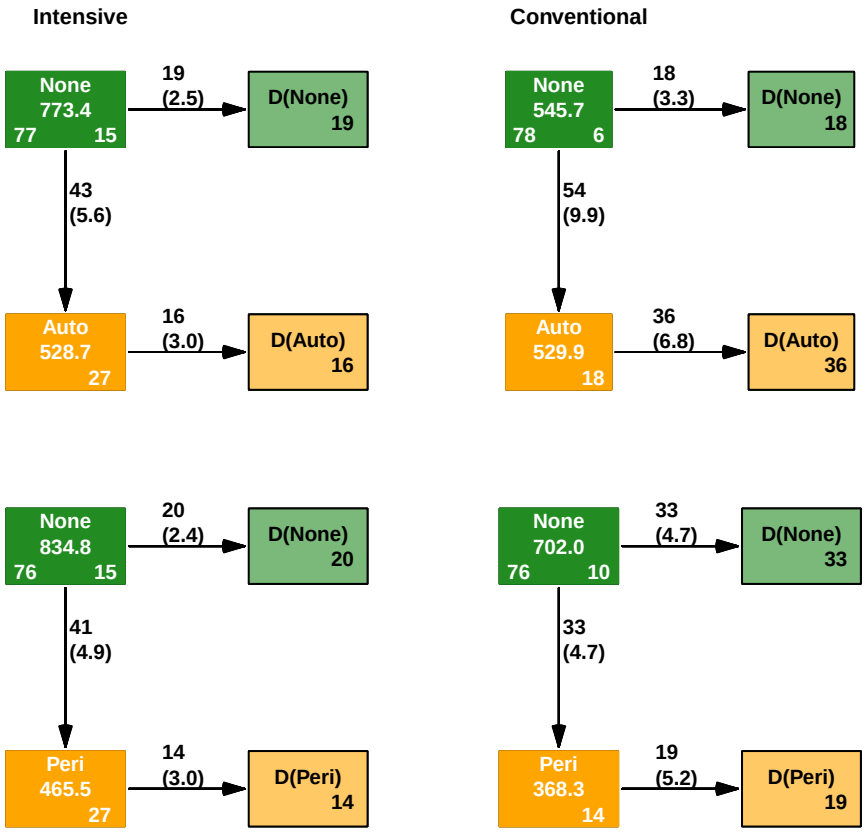

Figure 5.13: *Progression of autonomous and peripheral neuropthy, respectively. The number of persons is smaller that the total number randomized, because only persons with at least one assessment of the relevant type of neuropathy are included. The number in the center of the boxes is the person-years (PY), the numbers at the bottom is the number of patients starting, resp. ending their follow-up in each box (state), and the numbers on the arrows are the number of transitions and the overall transition rates per 100 PY.*

5.3.5 Autonomous neuropathy

Then we fit a simple Cox-model with allocation as the only covariate and amend it by age and sex:

```
> a1 <- coxph( Surv(tfe,tfe+lex.dur,lex.Xst=="Auto") ~ allocation, data = LxA)
> a2 <- update( a1, . ~ . + I((age-55)/10) + sex )
> round( ci.exp( a2 ), 3 )
```

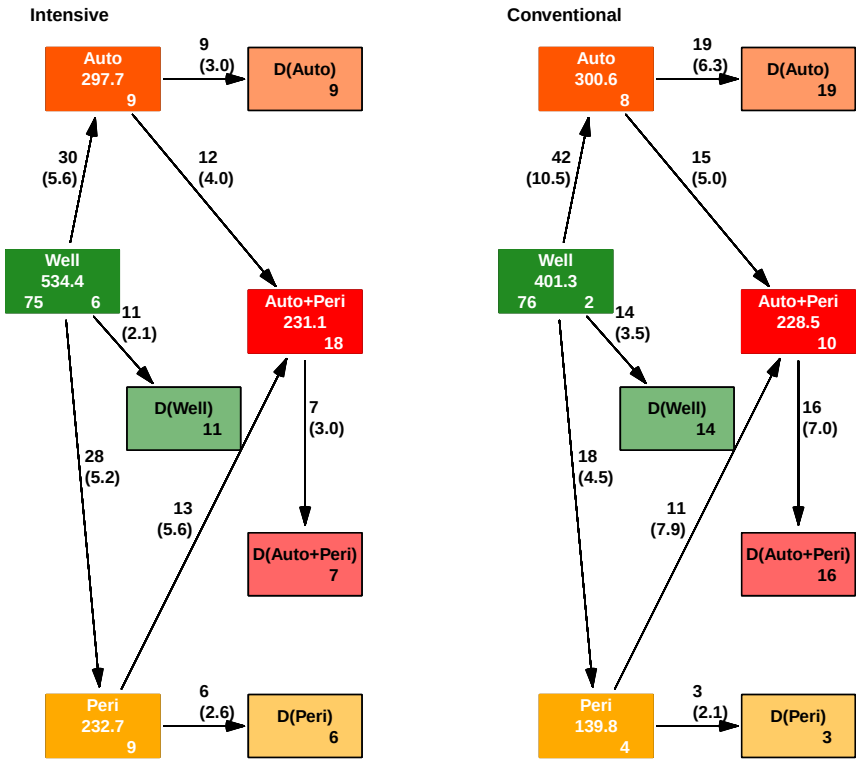

Figure 5.14: *Joint progression of autonomous and peripheral neuropthy. The number of persons is smaller that the total number randomized, because only persons with at least one assessment of both types of neuropathy are included. The number in the center of the boxes is the person-years (PY), the numbers at the bottom is the number of patients starting, resp. ending their follow-up in each box (state), and the numbers on the arrows are the number of transitions and the overall transition rates per 100 PY.*

```
                      exp(Est.)  2.5% 97.5%
allocationConventional     1.303 1.047 1.622
I((age - 55)/10)           1.020 0.870 1.196
sexMale                    1.299 1.000 1.687
> HR <- rbind( ci.exp( a1, pval=TRUE ),
+              ci.exp( a2, pval=TRUE ) )
> HR <- cbind( 1/HR[,c(1,3,2)], HR )
> rownames( HR )[1:2] <- c("Unadj.","Adj sex age")
> colnames( HR )[c(1,4)] <- c("Int vs Conv","Conv vs Int")
> round( HR[1:2,], 3 )
            Int vs Conv 97.5%  2.5% Conv vs Int  2.5% 97.5%     P
Unadj.            0.774 0.622 0.963       1.292 1.038 1.608 0.022
Adj sex age       0.767 0.616 0.955       1.303 1.047 1.622 0.018
```

So we see that there is a 25% lover occurrence rate of autonomous neuropathy in the intervention group (HR=0.75, 95% c.i.: 0.59–0.95), p=0.017, hence a significant effect of intervention, but the confidence interval covers both trivial and substantial effects. The model controlled for age and sex shows a similar intervention effect (HR 0.73 (0.57,0.93)), and reveals that there is virtually no age-effect but a 40% higher HR of autonomous neuropathy among men as compared to women.

5.3.5.1 Prediction of burden

In order to predict the risk of autonomous neuropathy we fit parametric models for the transitions. So we split the follow-up in pieces of 1 month:

```
> SxA <- splitLexis( LxA, breaks=seq(0,25,1/12), time.scale="tfe" )
```

For convenience we use one set of knots for the mortalities, and one set for incidence of neuropathy:

```
> summary( LxA )
Transitions:
     To
From   None Auto D(None) D(Auto)  Records:  Events: Risk time:  Persons:
  None  338   97      37       0       472      134    1319.06       155
  Auto    0  229       0      52       281       52    1058.67        97
  Sum   338  326      37      52       753      186    2377.73       155
> summary( SxA )
Transitions:
     To
From    None  Auto D(None) D(Auto)  Records:  Events: Risk time:  Persons:
  None 16091    97      37       0     16225      134    1319.06       155
  Auto     0 12932       0      52     12984       52    1058.67        97
  Sum  16091 13029      37      52     29209      186    2377.73       155
> ( d.kn <- with( subset( SxA, lex.Xst %in% c("D(None)","D(Auto)") ),
+                 quantile( tfe+lex.dur, probs=(1:4-0.5)/4 ) ) )
    12.5%     37.5%     62.5%     87.5%
 5.426420  9.119781 12.479124 16.870637
> ( n.kn <- with( subset( LxA, lex.Cst=="None" & lex.Xst=="Auto" ),
+                 quantile( tfe+lex.dur, probs=(1:4-0.5)/4 ) ) )
    12.5%     37.5%     62.5%     87.5%
 1.294970  3.442148  6.201048 10.236395
```

We fit models for mortality including age and sex and test for interaction between allocation and time since baseline:

```
> m1 <- glm( (lex.Xst %in% c("D(None)","D(Auto)")) ~
+            Ns( tfe, knots=d.kn) + sex + I((age-tfe)/10) + allocation + lex.Cst,
+            offset=log(lex.dur),
+            family=poisson,
+            data = SxA )
> mi <- update( m1, . ~ . + allocation:lex.Cst )
> mn <- update( m1, . ~ . + allocation:tfe )
> anova( mn, m1, mi, test="Chisq")
Analysis of Deviance Table

Model 1: (lex.Xst %in% c("D(None)", "D(Auto)")) ~ Ns(tfe, knots = d.kn) +
    sex + I((age - tfe)/10) + allocation + lex.Cst + allocation:tfe
Model 2: (lex.Xst %in% c("D(None)", "D(Auto)")) ~ Ns(tfe, knots = d.kn) +
    sex + I((age - tfe)/10) + allocation + lex.Cst
Model 3: (lex.Xst %in% c("D(None)", "D(Auto)")) ~ Ns(tfe, knots = d.kn) +
    sex + I((age - tfe)/10) + allocation + lex.Cst + allocation:lex.Cst
  Resid. Df Resid. Dev Df  Deviance Pr(>Chi)
1     29200     1130.7
2     29201     1130.7 -1 -0.006715   0.9347
3     29200     1130.4  1  0.277018   0.5987
> round( ci.exp( m1, pval=T ), 3 )
                        exp(Est.)  2.5%  97.5%     P
(Intercept)                 0.000 0.000  0.000 0.000
Ns(tfe, knots = d.kn)1      1.512 0.692  3.305 0.300
Ns(tfe, knots = d.kn)2      7.023 3.238 15.232 0.000
Ns(tfe, knots = d.kn)3      1.332 0.760  2.335 0.317
sexMale                     1.384 0.833  2.300 0.209
I((age - tfe)/10)           2.803 1.991  3.948 0.000
allocationConventional      2.076 1.339  3.220 0.001
lex.CstAuto                 0.969 0.619  1.517 0.890
```

We see that there is no sign of interaction between allocation and state (None/Auto) nor time since baseline — both p-values are over 0.7. Moreover there is no difference in mortality between those with or without autonomous neuropathy.

The occurrence of autonomous neuropathy is modelled likewise (this is essentially the same model as the Cox-model above, except that we model the baseline intensity parametrically):

```
> n1 <- glm( (lex.Xst == "Auto") ~
+            Ns( tfe, knots=d.kn) + sex + I(age-tfe) + allocation,
+            offset=log(lex.dur),
+            family=poisson,
+            data = subset(SxA, lex.Cst=="None" ) )
> ni <- update( n1, . ~ . + allocation:tfe )
> n2 <- update( ni, . ~ . + allocation:I(tfe^2) )
> aov <- anova( n1, ni, n2, test="Chisq")[-1,3:5]
> rownames( aov ) <- c("allo x time","allo x t-sq")
> round( aov, 3 )
            Df Deviance Pr(>Chi)
allo x time  1    0.964    0.326
allo x t-sq  2    5.728    0.057
> round( ci.exp( ni, pval=T ), 3 )
                           exp(Est.)  2.5% 97.5%     P
(Intercept)                    0.011 0.002 0.064 0.000
Ns(tfe, knots = d.kn)1         0.934 0.222 3.929 0.925
Ns(tfe, knots = d.kn)2         0.166 0.035 0.795 0.025
Ns(tfe, knots = d.kn)3         0.054 0.006 0.520 0.011
sexMale                        1.509 0.935 2.436 0.092
I(age - tfe)                   1.022 0.994 1.051 0.128
allocationConventional         2.344 1.135 4.841 0.021
allocationIntensive:tfe        1.056 0.946 1.180 0.332
allocationConventional:tfe     1.000 1.000 1.000   NaN
```

We then update the HR collector for autonomic neuropathy

```
> nz <- update( n1, . ~ . - I(age-tfe) - sex )
> zz <- rbind( ci.exp( nz, subset="allo", pval=TRUE ),
+              ci.exp( n1, subset="allo", pval=TRUE ) )
> zz[,1:3] <- 1/zz[,c(1,3,2)]
> load( file="../data/mainCI.Rda" )
> dimnames( mainCI )[[1]][8]
[1] "Autonomic neuropathy"
> mainCI[8,,] <- zz
> round( ftable(mainCI,row.vars=1), 3 )
                      model      raw                      age/sex
                      what  Estimate  2.5% 97.5%     P Estimate  2.5% 97.5%     P
outcome
All cause mortality             0.551 0.364 0.835 0.005    0.526 0.346 0.800 0.003
CVD mortality                   0.379 0.191 0.754 0.006    0.353 0.177 0.705 0.003
non-CVD mortality               0.703 0.413 1.197 0.195    0.686 0.401 1.172 0.168
Death or 1st CVD                   NA    NA    NA    NA       NA    NA    NA    NA
Death | CVD state               0.835 0.536 1.301 0.425    0.874 0.564 1.355 0.548
CVD event | CVD state           0.552 0.394 0.772 0.001    0.575 0.412 0.803 0.001
Retinopthy progression          0.668 0.507 0.881 0.004    0.673 0.511 0.887 0.005
Autonomic neuropathy            0.594 0.397 0.890 0.011    0.573 0.382 0.859 0.007
Peripheral neuropathy           1.120 0.707 1.774 0.630    1.101 0.694 1.747 0.683
Macroalbuminuria                0.516 0.316 0.842 0.008    0.495 0.302 0.811 0.005
> save( mainCI, file="../data/mainCI.Rda" )
```

There is a non-significant interaction with time since baseline, with decreasing allocation effect by time as seen in figure 5.15, and the curvature term is borderline significant. We can show how the estimated linear and quadratic effects look:

```
> tpt <- seq(2,20,0.5)
> CM  <- cbind(-1,tpt,-tpt,tpt^2,-(tpt^2))
> RRc <- ci.exp( n1, subset="allo", ctr.mat=CM[,1,drop=F] )
> RRl <- ci.exp( ni, subset="allo", ctr.mat=CM[,1:3] )
> RRq <- ci.exp( n2, subset="allo", ctr.mat=CM )
> matplot( tpt, cbind(RRc,RRq,RRl),
+          type="l", lty=1, lwd=c(3,1,1),
+          log="y", ylim=c(0.1,5),
+          col=rep(c(gray(0.5),"blue"),c(6,3)),
+          xlab="Time since baseline (years)",
+          ylab="HR of autonomous neuropathy: Int. vs Conv." )
> axis( side=1, at=2:20, tcl=-0.4, labels=NA )
> abline( h=1, col=gray(0.7) )
```

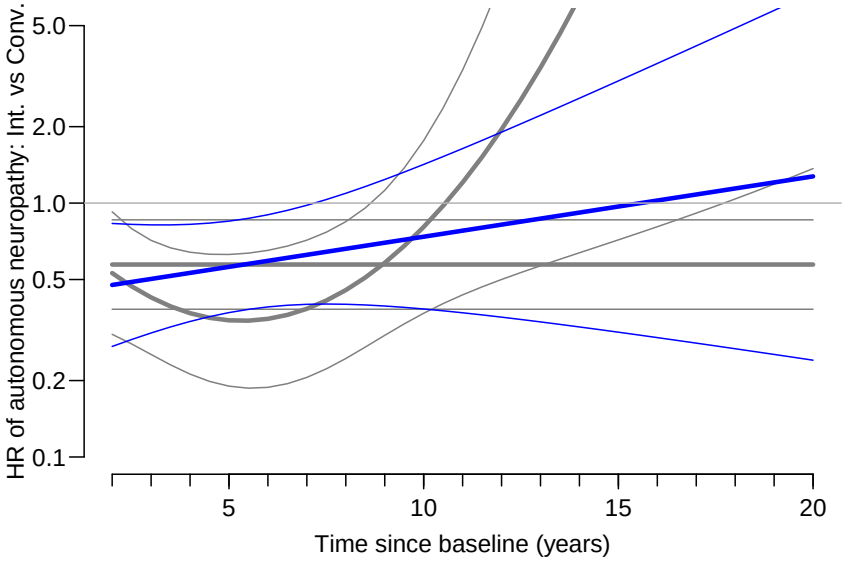

Figure 5.15: *Hazard ratio of autonomous neuropathy progression between intensive and conventional arms by time since baseline. Three different models: Constant (proportional hazards, gray), linear (blue) and quadratic (gray) interactions with time.*

The linear interaction model is not better that the the proportional hazards model (P==327), and althogh the curvature term is borderline signifcant (P=0.058), the shap of it as shown in figure 5.15 is not credible and so we shall stick to the proportional hazards model.

With these models fitted we shall show the probability of having autonomous neuropathy in the two groups, so we set up a transition object as well as a a baseline object from which to simulate.

```
> trA <- list( "None" = list(   "Auto"  = n1,
+                             "D(None)" = m1 ),
+              "Auto" = list( "D(Auto)" = m1 ) )
> lapply( trA, names )
$None
[1] "Auto"    "D(None)"

$Auto
[1] "D(Auto)"
```

Since autonomous neuropathy is only measured as from the 2nd visit, we set the starting time to 2 years, but we use as initiating state for each person the status at the first

measurement of autonomous neuropathy (which for a few persons is at a later visit than the 2nd). We use two copies of this dataset (that is the distribution of age, sex and neuropathy) one where we assume intensive treatment and one where we assume conventional:

```
> # selecting columns with "[]" loses the attributes so this is not applicable:
> # iniA <- LxA[match(unique(LxA$lex.id),LxA$lex.id),c("sex","allocation","lex.Cst","tfe","age","per
> iniA <- subset( LxA, per==ave(per,lex.id,FUN=min),
+                 select=c("sex","allocation","lex.Cst","tfe","age","per") )
> iniA$tfe = 0
> iniA <- rbind(iniA,iniA)
> iniA$allocation <- factor( rep( levels(iniA$allocation),
+                                 each=nrow(iniA)/2 ) )
> str( iniA )
Classes 'Lexis' and 'data.frame':       310 obs. of  6 variables:
 $ sex       : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 1 2 2 2 ...
 $ allocation: Factor w/ 2 levels "Conventional",..: 2 2 2 2 2 2 2 2 2 2 ...
 $ lex.Cst   : Factor w/ 4 levels "None","Auto",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ tfe       : num  0 0 0 0 0 0 0 0 0 0 ...
 $ age       : num  61.1 46.5 50 48.6 57.2 ...
 $ per       : num  1993 1993 1993 1993 1993 ...
 - attr(*, "breaks")=List of 3
  ..$ per: NULL
  ..$ age: NULL
  ..$ tfe: NULL
 - attr(*, "time.scales")= chr  "per" "age" "tfe"
 - attr(*, "time.since")= chr  "" "" ""
> with(iniA,table(allocation,lex.Cst))
              lex.Cst
allocation     None Auto D(None) D(Auto)
  Conventional  155    0       0       0
  Intensive     155    0       0       0
```

Then we can simulate the neuropathy development and death occurrence in a hypothetical cohort with the sex- and age-distribution as the entire steno 2 cohort.

```
> system.time( sA <- simLexis( Tr=trA, init=iniA, N=100, time.pts=seq(0,22,0.2) ) )
   user  system elapsed
 72.799   0.548  73.336
> save( sA, file="../data/sA.Rda" )


> load( file="../data/sA.Rda" )
> summary( sA )
Transitions:
     To
From   None  Auto D(None) D(Auto)  Records:  Events: Risk time:  Persons:
  None 4267 19355    7378       0     31000    26733   270061.4     31000
  Auto    0  8254       0   11101     19355    11101   220557.0     19355
  Sum  4267 27609    7378   11101     50355    37834   490618.5     31000
```

In order to show how persons fare with respect to development of autonomous neuropathy and death we compute the state distribution by time since entry, starting at 2 years:

```
> prm <- c(1,2,4,3)
> pAi <- pState( nState( subset(sA,allocation=="Intensive"),
+                       at=seq(0,20,0.1), from=0, time.scale="tfe" ), prm )
> pAc <- pState( nState( subset(sA,allocation=="Conventional"),
+                       at=seq(0,20,0.1), from=0, time.scale="tfe" ), prm )
```

We can then plot the state distribution from 2 years, assuming the initial state distribution as the observed in the entire steno2 population:

```
> neust <-
+ function( Pi, Pc, rv=FALSE, leg=rep(15,4), xl=c(0,20) )
+ {
+ par( mfrow=c(1,2), mar=c(0,3,0,0), oma=c(3,0,2,3-2*rv), las=1 )
+ plot( Pi, col=clr[prm], xlim=xl, xlab="", # yaxt="n",
+           col.lab="transparent" )
+ lines( as.numeric(rownames(Pi)), Pi[,2], lwd=4 )
+ Pi <- cbind(0,Pi)
+ wr <- cbind( match(paste(leg),rownames(Pi)), 2:5 )
+ wl <- cbind( match(paste(leg),rownames(Pi)), 1:4 )
+ if(is.numeric(leg)) text( leg,
+                           (Pi[wl]+Pi[wr])/2,
+                           colnames(Pi)[-1],
+                           col=rep(c("white","black"),each=2) )
+ axis( side=2, at=1:10/10, labels=NA )
+ axis( side=4, at=0:5/5 )
+ axis( side=4, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4, at=0:100  /100, tcl=-0.2, labels=NA )
+ mtext( "Probability", side=2, line=2, las=0 )
+ mtext( "Intensive", side=3, line=1 )
+
+ plot( Pc, col=clr[prm], xlim=if(rv) xl[2:1] else xl,
+       xlab="", yaxt="n", col.lab="transparent" )
+ lines( as.numeric(rownames(Pc)), Pc[,2], lwd=4 )
+ axis( side=2+2*rv, at=1:10/10, labels=NA )
+ axis( side=4-2*rv, at=1:5/5, if(rv) labels=NA )
+ axis( side=4-2*rv, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4-2*rv, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4-2*rv, at=0:100  /100, tcl=-0.2, labels=NA )
+ mtext( "Conventional", side=3, line=1 )
+ mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
+ }
> neust( pAi, pAc, rv=FALSE, xl=c(0,20) )


> neust( pAi, pAc, rv=TRUE, xl=c(0,20) )
```

5.3.6 Peripheral neuropathy

Then we fit a simple Cox-model for the progression of peripheral neuropathy with allocation as the only covariate and amend it by age and sex:

```
> summary( LxP )
Transitions:
     To
From   None Peri D(None) D(Peri)  Records:  Events: Risk time:  Persons:
  None  264   74      53       0       391      127    1536.82       152
  Peri    0  168       0      33       201       33     833.78        74
  Sum   264  242      53      33       592      160    2370.59       152
> p1 <- coxph( Surv(tfe,tfe+lex.dur,lex.Xst=="Peri") ~ allocation, data = LxP)
> p2 <- update( p1, . ~ . + I((age-55)/10) + sex )
> round( ci.exp( p2, pval=TRUE ), 3 )
                       exp(Est.)  2.5% 97.5%     P
allocationConventional     0.941 0.729 1.214 0.640
I((age - 55)/10)           0.795 0.663 0.954 0.014
sexMale                    0.937 0.707 1.241 0.650
> HR <- rbind( ci.exp( p1, pval=TRUE ),
+              ci.exp( p2, pval=TRUE ) )
> round( HR[3:4,], 3 )
                  exp(Est.)  2.5% 97.5%     P
I((age - 55)/10)      0.795 0.663 0.954 0.014
sexMale               0.937 0.707 1.241 0.650
```

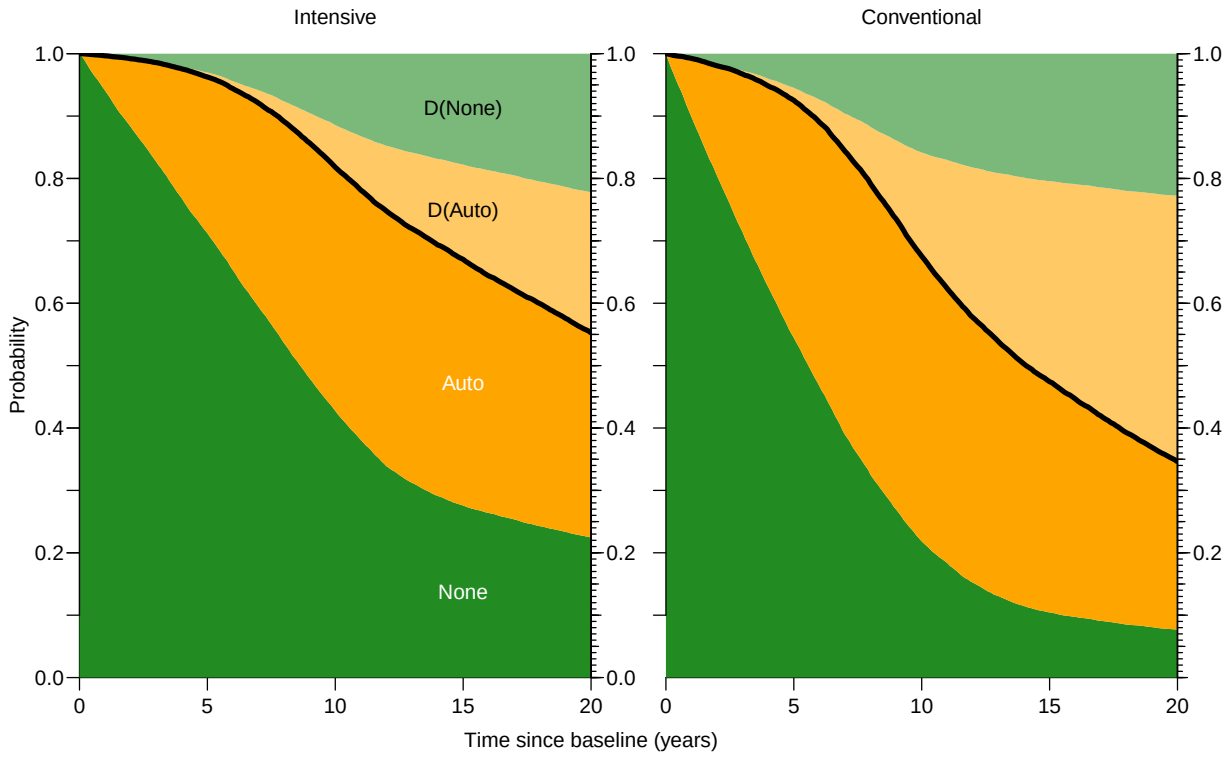

Figure 5.16: *Probabilities of autonomous neuropathy, using the baseline distribution of age, sex and autonomous neuropathy for the entire Steno 2 cohort.*

```
> HR <- cbind( 1/HR[,c(1,3,2)], HR )
> rownames( HR )[1:2] <- c("Unadj.","Adj sex age")
> colnames( HR )[c(1,4)] <- c("Int vs Conv","Conv vs Int")
> round( HR[1:2,], 3 )
            Int vs Conv 97.5%  2.5% Conv vs Int  2.5% 97.5%     P
Unadj.            1.063 0.824 1.372       0.941 0.729 1.214 0.639
Adj sex age       1.063 0.824 1.371       0.941 0.729 1.214 0.640
```

So we see that there is no intervention effect on peripheral neuropathy. The model controlled for age and sex shows the same intervention effect (HR 1.02 (0.79,1.33)), but reveals that there is a significantly decreasing effect of age, HR=0.79 per 10 years of age (95% c.i. 0.65–0.95, p=0.010).

5.3.6.1 Prediction of burden

In order to predict the risk of peripheral neuropathy we fit parametric models for the transitions. So we split the follow-up in pieces of 1 month:

```
> SxP <- splitLexis( LxP, breaks=seq(0,25,1/12), time.scale="tfe" )
```

For convenience we use one set of knots for the mortalities, and one set for incidence of neuropathy:

```
> summary( LxP )
Transitions:
     To
From   None Peri D(None) D(Peri)  Records:  Events: Risk time:  Persons:
  None  264   74      53       0       391      127    1536.82       152
  Peri    0  168       0      33       201       33     833.78        74
  Sum   264  242      53      33       592      160    2370.59       152
```

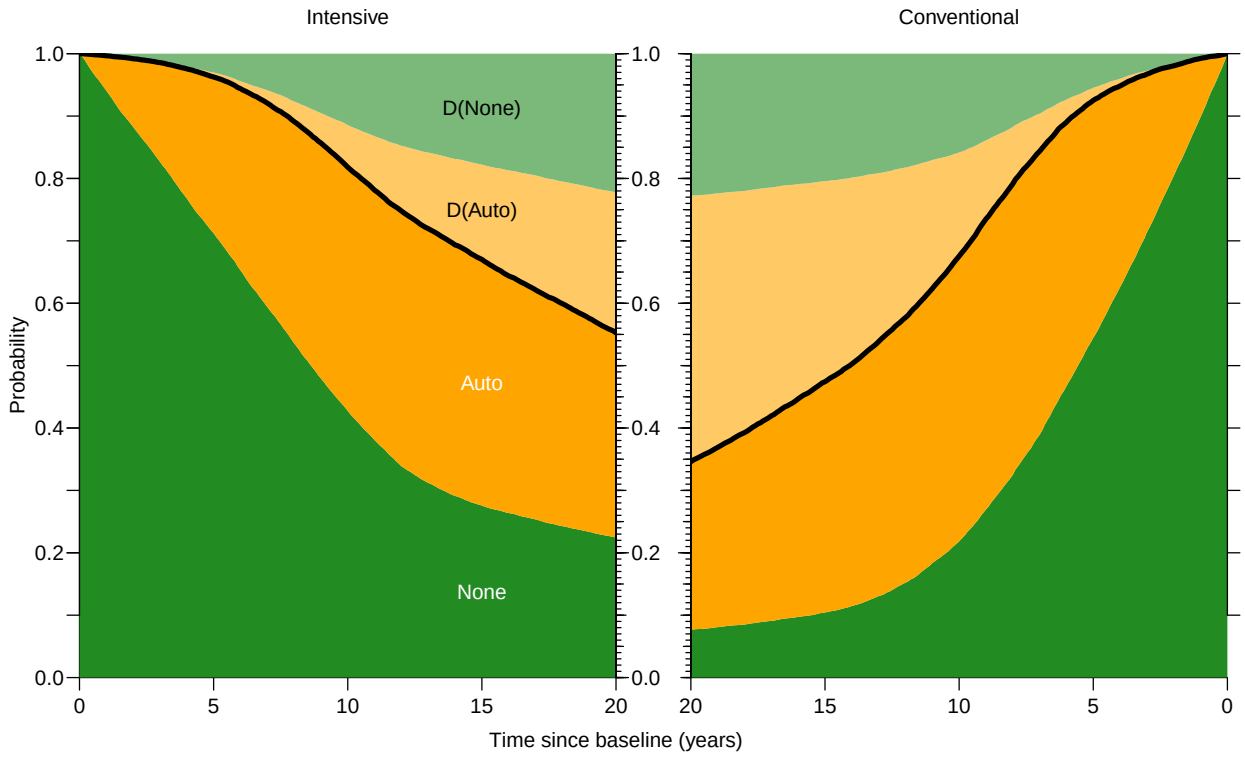

Figure 5.17: *Probabilities of autonomous neuropathy, using the baseline distribution of age, sex and autonomous neuropathy for the entire Steno 2 cohort. This is the same plot as 5.16, except that the plot for the conventional group is mirrored.*

```
> summary( SxP )

Transitions:
     To
From    None  Peri D(None) D(Peri)  Records:  Events: Risk time:  Persons:
  None 18627    74      53       0     18754      127    1536.82       152
  Peri     0 10177       0      33     10210       33     833.78        74
  Sum  18627 10251      53      33     28964      160    2370.59       152

> ( d.kn <- with( subset( SxP, lex.Xst %in% c("D(None)","D(Peri)") ),
+                 quantile( tfe+lex.dur, probs=(1:4-0.5)/4 ) ) )

    12.5%     37.5%     62.5%     87.5%
 6.061602  9.202601 12.645106 16.907598

> ( n.kn <- with( subset( LxP, lex.Cst=="None" & lex.Xst=="Peri" ),
+                 quantile( tfe+lex.dur, probs=(1:4-0.5)/4 ) ) )

    12.5%     37.5%     62.5%     87.5%
 1.091433  3.637061  6.096798 10.651029
```

We fit models for mortality including age and sex and test for interaction between allocation and time since baseline:

```
> m1 <- glm( (lex.Xst %in% c("D(None)","D(Peri)")) ~
+            Ns( tfe, knots=d.kn) + sex + I((age-tfe)/10) + allocation + lex.Cst,
+            offset=log(lex.dur),
+            family=poisson,
+            data = SxP )
> mi <- update( m1, . ~ . + allocation:lex.Cst )
> mn <- update( m1, . ~ . + allocation:tfe )
> anova( mn, m1, mi, test="Chisq")
```

```
Analysis of Deviance Table

Model 1: (lex.Xst %in% c("D(None)", "D(Peri)")) ~ Ns(tfe, knots = d.kn) +
    sex + I((age - tfe)/10) + allocation + lex.Cst + allocation:tfe
Model 2: (lex.Xst %in% c("D(None)", "D(Peri)")) ~ Ns(tfe, knots = d.kn) +
    sex + I((age - tfe)/10) + allocation + lex.Cst
Model 3: (lex.Xst %in% c("D(None)", "D(Peri)")) ~ Ns(tfe, knots = d.kn) +
    sex + I((age - tfe)/10) + allocation + lex.Cst + allocation:lex.Cst
  Resid. Df Resid. Dev Df  Deviance Pr(>Chi)
1     28955     1091.5
2     28956     1091.5 -1 -0.020599   0.8859
3     28955     1091.3  1  0.237259   0.6262
```

```
> round( ci.exp( m1, pval=T ), 3 )
```

```
                        exp(Est.)  2.5%  97.5%     P
(Intercept)                 0.000 0.000  0.000 0.000
Ns(tfe, knots = d.kn)1      1.334 0.601  2.962 0.479
Ns(tfe, knots = d.kn)2      6.810 3.233 14.348 0.000
Ns(tfe, knots = d.kn)3      1.183 0.670  2.086 0.563
sexMale                     1.385 0.828  2.315 0.214
I((age - tfe)/10)           2.712 1.910  3.852 0.000
allocationConventional      2.057 1.326  3.191 0.001
lex.CstPeri                 0.916 0.582  1.441 0.704
```

We see that there is no sign of interaction between allocation and state (None/Peri) nor time since baseline — both p-values are larger then 0.5. Moreover, there is no difference in mortality between those with and without progression of peripheral neuropathy (HR 0.87 (0.55;1.38), P=0.557).

The occurrence of peripheral neuropathy is modelled likewise (this is essentially the same model as the Cox-model above, except that we model the baseline intensity parametrically):

```
> n1 <- glm( (lex.Xst == "Peri") ~
+            Ns( tfe, knots=d.kn) + sex + I(age-tfe) + allocation,
+            offset=log(lex.dur),
+            family=poisson,
+            data = subset(SxP, lex.Cst=="None" ) )
> ni <- update( n1, . ~ . + allocation:tfe )
> n2 <- update( ni, . ~ . + allocation:I(tfe^2) )
> aov <- anova( n1, ni, n2, test="Chisq")[-1,3:5]
> rownames( aov ) <- c("allo x time","allo x t-sq")
> round( aov, 3 )
```

```
            Df Deviance Pr(>Chi)
allo x time  1    1.255    0.263
allo x t-sq  2    1.654    0.437
```

```
> round( ci.exp( n1, pval=T ), 3 )
```

```
                        exp(Est.)  2.5% 97.5%     P
(Intercept)                 0.168 0.027 1.054 0.057
Ns(tfe, knots = d.kn)1      0.289 0.073 1.153 0.079
Ns(tfe, knots = d.kn)2      0.341 0.151 0.771 0.010
Ns(tfe, knots = d.kn)3      0.333 0.133 0.833 0.019
sexMale                     0.950 0.569 1.589 0.846
I(age - tfe)                0.983 0.951 1.016 0.301
allocationConventional      0.908 0.572 1.442 0.683
```

There are no significant interactions with time since baseline, so a proportional hazards model seems to be reasonable to apply in prediction.

We then update the HR collector for peripheral neuropathy

```
> nz <- update( n1, . ~ . - I(age-tfe) - sex )
> zz <- rbind( ci.exp( nz, subset="allo", pval=TRUE ),
+              ci.exp( n1, subset="allo", pval=TRUE ) )
> zz[,1:3] <- 1/zz[,c(1,3,2)]
> load( file="../data/mainCI.Rda" )
> dimnames( mainCI )[[1]][9]
[1] "Peripheral neuropathy"
> mainCI[9,,] <- zz
> round( ftable(mainCI,row.vars=1), 3 )
                      model      raw                     age/sex
                      what  Estimate  2.5% 97.5%     P Estimate  2.5% 97.5%     P
outcome
All cause mortality             0.551 0.364 0.835 0.005    0.526 0.346 0.800 0.003
CVD mortality                   0.379 0.191 0.754 0.006    0.353 0.177 0.705 0.003
non-CVD mortality               0.703 0.413 1.197 0.195    0.686 0.401 1.172 0.168
Death or 1st CVD                   NA    NA    NA    NA       NA    NA    NA    NA
Death | CVD state               0.835 0.536 1.301 0.425    0.874 0.564 1.355 0.548
CVD event | CVD state           0.552 0.394 0.772 0.001    0.575 0.412 0.803 0.001
Retinopthy progression          0.668 0.507 0.881 0.004    0.673 0.511 0.887 0.005
Autonomic neuropathy            0.594 0.397 0.890 0.011    0.573 0.382 0.859 0.007
Peripheral neuropathy           1.120 0.707 1.774 0.630    1.101 0.694 1.747 0.683
Macroalbuminuria                0.516 0.316 0.842 0.008    0.495 0.302 0.811 0.005
> save( mainCI, file="../data/mainCI.Rda" )
```

With these models fitted we shall show the probability of having peripheral neuropathy in the two groups, so we set up a transition object as well as a a baseline object from which to simulate.

```
> trP <- list( "None" = list(   "Peri"  = n1,
+                             "D(None)" = m1 ),
+              "Peri" = list( "D(Peri)" = m1 ) )
> lapply( trP, names )
$None
[1] "Peri"    "D(None)"

$Peri
[1] "D(Peri)"
```

Since peripheral neuropathy is only measured as from the 3rd visit, we set the starting time to 4 years, but we use as initiating state for each person the status at the first measurement of peripheral neuropathy (which for a few persons is at a later visit than the 2nd). We use two copies of this dataset (that is the distribution of age, sex and neuropathy) one where we assume intensive treatment and one where we assume conventional:

```
> # selecting columns with "[]" loses the attributes so this is not applicable:
> # iniP <- LxP[match(unique(LxP$lex.id),LxP$lex.id),c("sex","allocation","lex.Cst","tfe","age","per
> iniP <- subset( LxP, per==ave(per,lex.id,FUN=min),
+                 select=c("sex","allocation","lex.Cst","tfe","age","per") )
> iniP$tfe = 0
> iniP <- rbind(iniP,iniP)
> iniP$allocation <- factor( rep( levels(iniP$allocation),
+                                 each=nrow(iniP)/2 ) )
> str( iniP )
Classes 'Lexis' and 'data.frame':       304 obs. of  6 variables:
 $ sex       : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 1 2 2 2 ...
 $ allocation: Factor w/ 2 levels "Conventional",..: 2 2 2 2 2 2 2 2 2 2 ...
 $ lex.Cst   : Factor w/ 4 levels "None","Peri",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ tfe       : num  0 0 0 0 0 0 0 0 0 0 ...
 $ age       : num  61.1 46.5 50 48.6 57.2 ...
 $ per       : num  1993 1993 1993 1993 1993 ...
 - attr(*, "breaks")=List of 3
  ..$ per: NULL
```

```
  ..$ age: NULL
  ..$ tfe: NULL
 - attr(*, "time.scales")= chr  "per" "age" "tfe"
 - attr(*, "time.since")= chr  "" "" ""
> with(iniP,table(allocation,lex.Cst))
              lex.Cst
allocation     None Peri D(None) D(Peri)
  Conventional  152    0       0       0
  Intensive     152    0       0       0
```

Then we can simulate the neuropathy development and death occurrence in a hypothetical cohort with the sex-, age- and neuropathy-distribution as the entire steno 2 cohort.

```
> system.time( sP <- simLexis( Tr=trP, init=iniP, N=100, time.pts=seq(0,22,0.2) ) )
   user  system elapsed
 62.831   0.336  63.157
> save( sP, file="../data/sP.Rda" )


> load( file="../data/sP.Rda" )
> summary( sP )
Transitions:
     To
From   None  Peri D(None) D(Peri)  Records:  Events: Risk time:  Persons:
  None 4898 14945   10557       0     30400    25502   314147.6     30400
  Peri    0  7737       0    7208     14945     7208   175101.0     14945
  Sum  4898 22682   10557    7208     45345    32710   489248.6     30400
```

In order to show how persons fare with respect to development of autonomous neuropathy and death we compute the state distribution by time since entry, starting at 2 years:

```
> prm <- c(1,2,4,3)
> pPi <- pState( nState( subset(sP,allocation=="Intensive"),
+                       at=seq(0,20,0.1), from=0, time.scale="tfe" ), prm )
> pPc <- pState( nState( subset(sP,allocation=="Conventional"),
+                       at=seq(0,20,0.1), from=0, time.scale="tfe" ), prm )
```

We can then plot the state distribution from 2 years, assuming the initial state distribution as the observed in the entire steno2 population:

```
> neust( pPi, pPc, rv=FALSE, xl=c(0,20) )


> neust( pPi, pPc, rv=TRUE, xl=c(0,20) )
```

From figure 5.19 we see that the intensive group have a 20-year survival proability of 56.6% out of which 33.7/ 56.6= 59.5% have peripheral neuropathy, whereas the conventional group have a 20-year survival proability of 35.2% out of which 12.2/ 35.2= 57.9% have peripheral neuropathy.

So largely the same fraction of the survivors at 20 years in the two groups.

Finally, we save the state-distributions for autonomous and peripheral neuropathy

```
> save( pAi, pAc, pPi, pPc, file="../data/pNeu.Rda" )
```

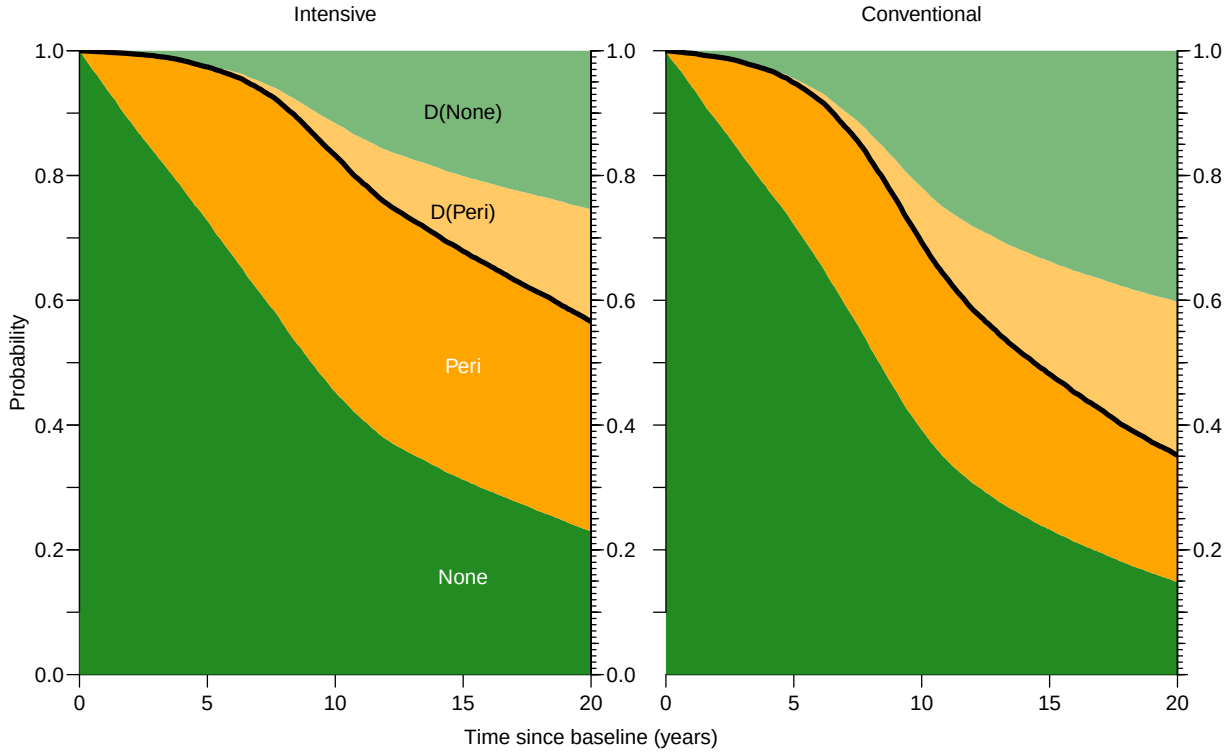

Figure 5.18: *Probabilities of peripheral neuropathy, using the baseline distribution of age, sex and peripheral neuropathy for the entire Steno 2 cohort. Note that the timescale starts at 4 years after baseline.*

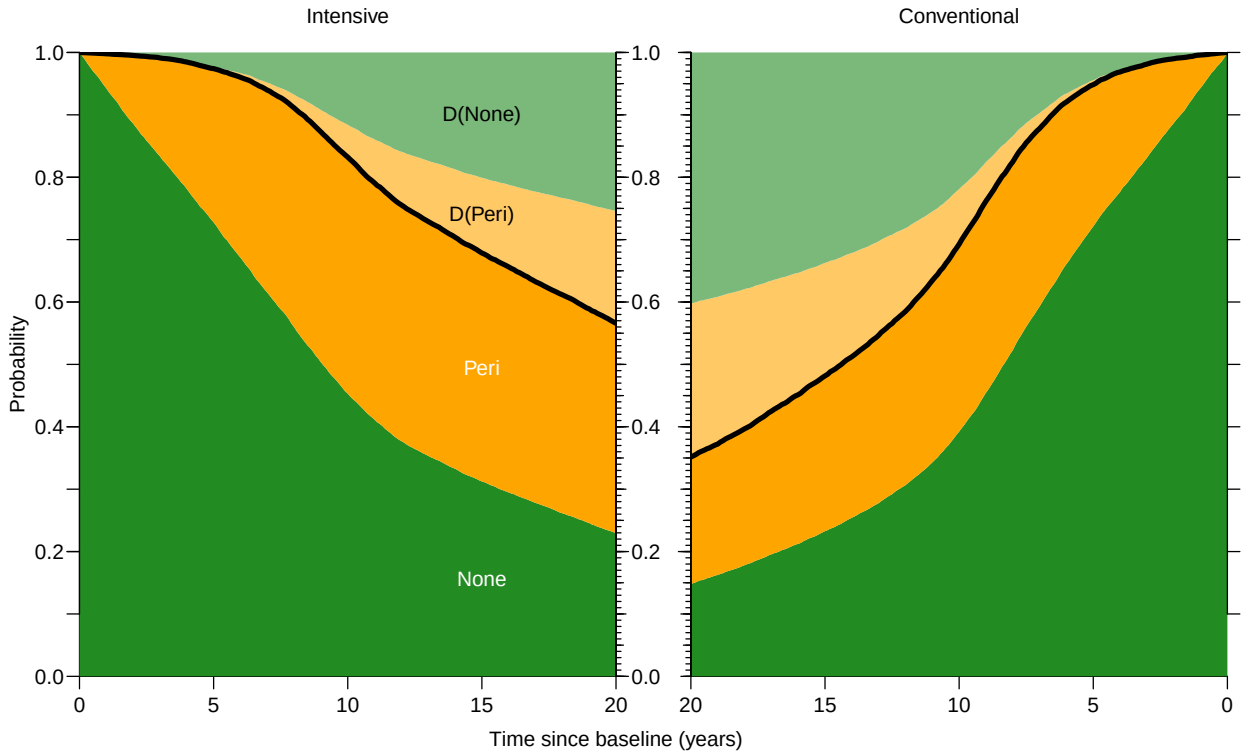

Figure 5.19: *Probabilities of peripheral neuropathy, using the baseline distribution of age, sex and peripheral neuropathy for the entire Steno 2 cohort. Note that the timescale starts at 4 years after baseline. This is the same plot as 5.18, except that the plot for the conventional group is mirrored.*

5.4 Albuminuria

5.4.1 Albumin measurements

First we read the data:

```
> library( Epi )
> library( survival )
> clear()
> load( file="../data/mic.Rda" )
> names( mic )
 [1] "ID"          "allocation"  "sex"         "doBth"       "doVis"       "doESRD"
 [7] "doEye"       "doPhC"       "doBlind"     "doEnd"       "doDth"       "visit"
[13] "aer_level"   "med_aer"     "eds_l"       "eds_r"       "autoprog"    "periprog"
```

We will need the time since first visit in order to get the variation between persons at baseline as a parameter:

```
> fv <- subset( mic, visit==1, select=c("ID","doVis") )
> names(fv)[2] <- "doBase"
> mic <- transform( merge( mic, fv, all=TRUE ),
+                   tfBase = doVis - doBase )
```

We then fit a model for the log of the median albumin measurement, allowing separate slopes in the two groups:

```
> library( lme4 )
> m1 <- lmer( log(med_aer) ~ allocation + allocation:tfBase + (tfBase|ID), data=mic )
> summary( m1 )
Linear mixed model fit by REML ['lmerMod']
Formula: log(med_aer) ~ allocation + allocation:tfBase + (tfBase | ID)
   Data: mic

REML criterion at convergence: 2204

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.4592 -0.5665  0.0109  0.5468  3.2020

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ID       (Intercept) 0.338363 0.5817
          tfBase      0.008483 0.0921   0.27
 Residual             0.766796 0.8757
Number of obs: 723, groups:  ID, 160

Fixed effects:
                            Estimate Std. Error t value
(Intercept)                  4.14191    0.09189   45.07
allocationConventional       0.21263    0.13157    1.62
allocationIntensive:tfBase   0.02347    0.01419    1.65
allocationConventional:tfBase  0.05305    0.01634    3.25

Correlation of Fixed Effects:
            (Intr) allctC allI:B
allctnCnvnt -0.698
allctnInt:B -0.169  0.118
allctnCnv:B  0.000 -0.177  0.000
```

We can make a formal test of equality of the two groups:

```
> round( Wald( m1, subset="allo", ctr.mat=rbind(c(1,0,0),c(0,1,-1)) ), 3 )
Chisq  d.f.     P
5.664 2.000 0.059
```

so in this model with a linear trend in log(AER) there is a borderline significant difference between the groups:

```
> round( cbind( ci.exp( m1 ), 1/ci.exp( m1 ) ), 3 )
                              exp(Est.)   2.5%  97.5% exp(Est.)  2.5% 97.5%
(Intercept)                      62.923 52.552 75.341     0.016 0.019 0.013
allocationConventional            1.237  0.956  1.601     0.808 1.046 0.625
allocationIntensive:tfBase        1.024  0.996  1.053     0.977 1.004 0.950
allocationConventional:tfBase     1.054  1.021  1.089     0.948 0.979 0.918
```

This means that the *ratio* of the median albumin levels between intensive and conventional at baseline is 0.81, and since the random effect for `ID` refer to the baseline the standard deviation of this (0.582) corresponds to the median difference in log-albumin levels between two randomly chosen persons. Specifically, it must be multiplied by $\sqrt{2}z_{0.75}$ and subsequently exponentiated to give the median ratio between two randomly chosen persons:

```
> str( VarCorr( m1 ) )
List of 1
 $ ID: num [1:2, 1:2] 0.33836 0.01471 0.01471 0.00848
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : chr [1:2] "(Intercept)" "tfBase"
  .. ..$ : chr [1:2] "(Intercept)" "tfBase"
  ..- attr(*, "stddev")= Named num [1:2] 0.5817 0.0921
  .. ..- attr(*, "names")= chr [1:2] "(Intercept)" "tfBase"
  ..- attr(*, "correlation")= num [1:2, 1:2] 1 0.275 0.275 1
  .. ..- attr(*, "dimnames")=List of 2
  .. .. ..$ : chr [1:2] "(Intercept)" "tfBase"
  .. .. ..$ : chr [1:2] "(Intercept)" "tfBase"
 - attr(*, "sc")= num 0.876
 - attr(*, "useSc")= logi TRUE
 - attr(*, "class")= chr "VarCorr.merMod"
> ( btw <- exp( sqrt(2)*qnorm(0.75)*attr(VarCorr(m1)$ID,"stddev")[1] ) )
(Intercept)
   1.741693
```

The number `\Sexpr[]{}`round(btw,2)) describes the between-person variation in the entire study population at baseline, so the baseline difference in albumin levels is quite small relative to the population variation — which of course is what should be expected in a clinical trial.

The estimated ratio of slopes in the two groups are:

```
> rt <- ci.exp( m1, subset="tfB", ctr.mat=rbind(c(-1,1)), pval=T )
> round( cbind( rt, 1/rt[,-4,drop=F] ), 3 )
     exp(Est.)  2.5% 97.5%     P exp(Est.)  2.5% 97.5%
[1,]      1.03 0.987 1.075 0.172     0.971 1.013 0.931
```

As before the between-person-variation in slopes can be extracted:

```
> exp( sqrt(2) * qnorm(0.75) * attr(VarCorr(m1)$ID,"stddev")[2] )
  tfBase
1.091831
```

so the median ratio in slopes between two randomly chosen persons is 1.09; substantially larger than the estimated ratio of the mean slopes.

Thus we conclude that there is little difference in the levels and in the rate of change in albumin levels between the intensive and conventional arm. If we want to assess the ratio of the AER levels between intensive and conventional groups at different times after baseline, we compute these by `ci.exp`; and we compute the between person variation by pre-post multiplying the contrast matrix:

```
> dim( VarCorr(m1) )
NULL
> sqrt( diag(VarCorr(m1)$ID) )
(Intercept)      tfBase
 0.58168974  0.09210464
> t.pt <- 0:20
> CM <- cbind( -1, t.pt, -t.pt )
> MC <- -CM[,-2]
> vv <- sqrt(diag(MC %*% VarCorr(m1)$ID %*% t(MC)))
> vv <- exp( sqrt(2) * qnorm(0.75) * vv )
> zz <- ci.exp( m1, subset="allocation", ctr.mat=CM, pval=T )
> matplot( t.pt, zz[,-4], log="y", ylim=c(0.15,1.5),
+          type="l", col="black", lty=1, lwd=c(3,1,1), las=1,
+          xlab="Time since baseline (years)",
+          ylab="AER ratio, Intensive vs Conventional" )
> axis( side=2, at=c(0.15,2:15/10), labels=NA )
> polygon( c(rev(t.pt),t.pt),
+          c(rep(1,length(t.pt)),1/vv),
+          col=rgb(180,180,180,50,max=255), border="transparent" )
> abline( h=1, lty=3 )
```

From figure 5.20 we see that there is a borderline significant improvement in the mean AER-rate, but also that this is small compared to the between person variation in AER rates. The AER-ratio between the intensive and the conventional group changes from 0.81 at baseline to 0.45 at 20 years.

5.4.1.1 0-intercept models

It might be more realistic, since we have a randomized study, to assume a priori that the ratio of AER between groups is 1 at baseline, and then allow for a curved relationship in the ratio:

```
> m1 <- lmer( log(med_aer) ~ allocation:tfBase +
+                            (tfBase|ID), data=mic )
> mq <- lmer( log(med_aer) ~ allocation:tfBase + allocation:I(tfBase^2) +
+                            (tfBase|ID), data=mic )
> summary( m1 )
Linear mixed model fit by REML ['lmerMod']
Formula: log(med_aer) ~ allocation:tfBase + (tfBase | ID)
   Data: mic

REML criterion at convergence: 2204.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.4576 -0.5560  0.0277  0.5477  3.2420

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ID       (Intercept) 0.347706 0.58967
          tfBase      0.008664 0.09308  0.25
 Residual             0.764292 0.87424
Number of obs: 723, groups:  ID, 160

Fixed effects:
                            Estimate Std. Error t value
(Intercept)                  4.24577    0.06618   64.16
allocationIntensive:tfBase   0.02053    0.01418    1.45
allocationConventional:tfBase  0.05789    0.01616    3.58

Correlation of Fixed Effects:
            (Intr) allI:B
allctnInt:B -0.125
allctnCnv:B -0.178  0.022
```

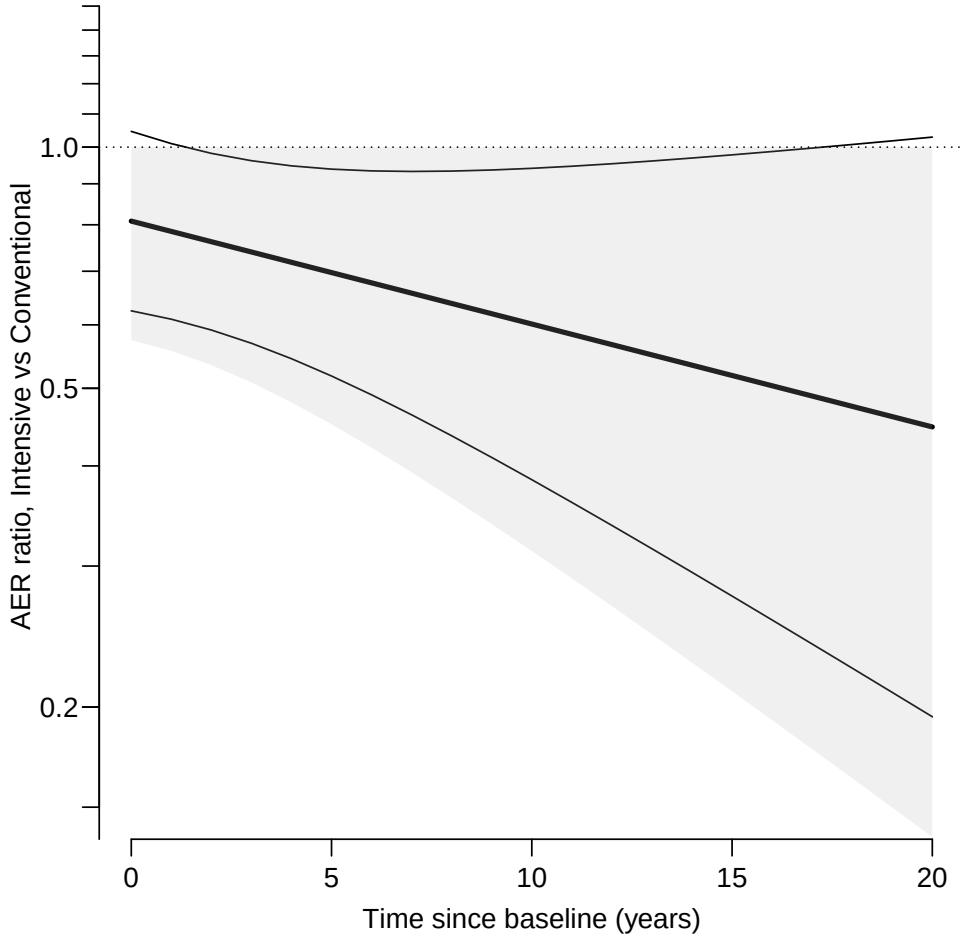

Figure 5.20: *Estimated mean AER-ratio between intensive and conventional by time since baseline. The gray area represents the between-person-variation: The height of the area is the median ratio of AER between two randomly selected persons from the same allocation group.*

```
> summary( mq )
Linear mixed model fit by REML ['lmerMod']
Formula: log(med_aer) ~ allocation:tfBase + allocation:I(tfBase^2) + (tfBase |      ID)
   Data: mic

REML criterion at convergence: 2206.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.5544 -0.5387 -0.0281  0.5169  3.2626

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ID       (Intercept) 0.337027 0.58054
          tfBase      0.009721 0.09859  0.25
 Residual             0.730010 0.85441
Number of obs: 723, groups:  ID, 160

Fixed effects:
                                Estimate Std. Error t value
(Intercept)                     4.343379   0.071214   60.99
```

```
allocationIntensive:tfBase         -0.071465   0.024769   -2.89
allocationConventional:tfBase       0.049159   0.027528    1.79
allocationIntensive:I(tfBase^2)     0.005491   0.001190    4.61
allocationConventional:I(tfBase^2)  0.000420   0.001577    0.27

Correlation of Fixed Effects:
            (Intr) allI:B allC:B aI:I(B
allctnInt:B -0.320
allctnCnv:B -0.336  0.107
allI:I(B^2)  0.311 -0.807 -0.105
allC:I(B^2)  0.292 -0.093 -0.799  0.091
```

In the model with linear trend in log-AER the annual changes and their difference are:

```
> ( CC <- rbind( Int = c(1,0),
+               Conv = c(0,1),
+              "I-C" = c(1,-1) ) )
     [,1] [,2]
Int     1    0
Conv    0    1
I-C     1   -1
> ci.exp( ml, subset="tfB" )
                               exp(Est.)      2.5%    97.5%
allocationIntensive:tfBase     1.020743 0.9927635 1.049510
allocationConventional:tfBase  1.059602 1.0265692 1.093698
> aer.1 <- ci.exp( ml, subset="tfB", ctr.mat=CC, pval=TRUE )
> cbind( round( (aer.1[,-4]-1)*100, 1 ),
+        round( aer.1, 3 ) )
     exp(Est.) 2.5% 97.5% exp(Est.)  2.5% 97.5%     P
Int        2.1 -0.7   5.0     1.021 0.993 1.050 0.148
Conv       6.0  2.7   9.4     1.060 1.027 1.094 0.000
I-C       -3.7 -7.6   0.4     0.963 0.924 1.004 0.079
```

Thus we see that there is a borderline significantly smaller decrease in the intervention group — 4% less per year,

A formal test of equality of the two groups in the liner 0-intercept model is:

```
> round( Wald( ml, subset="allocation", ctr.mat=rbind(c(-1,1)) ), 3 )
Chisq  d.f.     P
3.089 1.000 0.079
```

— so there is some (weak) evidence for a difference between the groups. We can compute the estimated ratio of means in the two groups as for the intercept model:

```
> CM <- cbind( t.pt, -t.pt )
> round( ll <- ci.exp( ml, subset="allocation", ctr.mat=CM, pval=T ), 3 )
      exp(Est.)  2.5% 97.5%     P
 [1,]     1.000 1.000 1.000   NaN
 [2,]     0.963 0.924 1.004 0.079
 [3,]     0.928 0.854 1.009 0.079
 [4,]     0.894 0.789 1.013 0.079
 [5,]     0.861 0.729 1.017 0.079
 [6,]     0.830 0.674 1.022 0.079
 [7,]     0.799 0.622 1.026 0.079
 [8,]     0.770 0.575 1.031 0.079
 [9,]     0.742 0.531 1.035 0.079
[10,]     0.714 0.491 1.040 0.079
[11,]     0.688 0.454 1.044 0.079
[12,]     0.663 0.419 1.048 0.079
[13,]     0.639 0.387 1.053 0.079
[14,]     0.615 0.358 1.058 0.079
[15,]     0.593 0.331 1.062 0.079
[16,]     0.571 0.306 1.067 0.079
```

```
[17,]     0.550 0.282 1.071 0.079
[18,]     0.530 0.261 1.076 0.079
[19,]     0.510 0.241 1.081 0.079
[20,]     0.492 0.223 1.085 0.079
[21,]     0.474 0.206 1.090 0.079
```

We can make an approximate Wald test for the quadratic extension (ignoring the random part of the model)

```
> round( ci.lin( mq, subset="2" ), 3 )
                                  Estimate StdErr     z    P   2.5% 97.5%
allocationIntensive:I(tfBase^2)      0.005  0.001 4.613 0.00  0.003 0.008
allocationConventional:I(tfBase^2)   0.000  0.002 0.266 0.79 -0.003 0.004
> round( Wald( mq, subset="2", ctr.mat=diag(2) ), 3 )
 Chisq   d.f.      P
21.303  2.000  0.000
```

A formal test of equality of the two groups in the quadratic model is:

```
> round( Wald( mq, subset="allocation", ctr.mat=rbind(c(-1,1,0,0),c(0,0,-1,1)) ), 3 )
 Chisq   d.f.      P
11.880  2.000  0.003
```

We can then compute the ratio of averages between the two groups in the quadratic model and plot this together with the ratio from the 0-intercept linear model:

```
> ( sqrt( diag( vc <- VarCorr(mq)$ID ) ) )
(Intercept)      tfBase
  0.5805400   0.0985947
> qv <- exp( sqrt(2) * qnorm(0.75) * sqrt(diag(MC %*% vc %*% t(MC))) )
> ci.exp( mq )
                                   exp(Est.)       2.5%      97.5%
(Intercept)                       76.9671487 66.9403723 88.4958027
allocationIntensive:tfBase         0.9310284  0.8869107  0.9773407
allocationConventional:tfBase      1.0503877  0.9952176  1.1086160
allocationIntensive:I(tfBase^2)    1.0055059  1.0031628  1.0078543
allocationConventional:I(tfBase^2) 1.0004201  0.9973317  1.0035180
> CM <- cbind( t.pt, -t.pt, t.pt^2, -(t.pt^2) )
> qq <- ci.exp( mq, subset="allocation", ctr.mat=CM, pval=T )
> matplot( t.pt, cbind( qq[,-4], ll[,-4] ), log="y", ylim=c(0.15,1.5),
+          type="l", col=rep(c(gray(0.6),"black"),each=3), lty=1, lwd=c(4,1,1), las=1,
+          xlab="Time since baseline (years)",
+          ylab="AER ratio, Intensive vs Conventional" )
> axis( side=2, at=c(0.15,2:15/10), labels=NA )
> polygon( c(rev(t.pt),t.pt),
+          c(rep(1,length(t.pt)),1/qv),
+          col=rgb(180,180,180,50,max=255), border="transparent" )
> abline( h=1, lty=3 )
```

From figure 5.21 it is clear that the intervention effect is borderline significant; formally there is a significant difference between the groups (p=0.003), and the average ratio of AER levels between intensive and conventional allocation is about 0.4.

5.4.2 Normo-, micro- and macro-albuminuria

We load the micro-vascular data; that is records from each follow-up visit, s well as the functions needed to generate

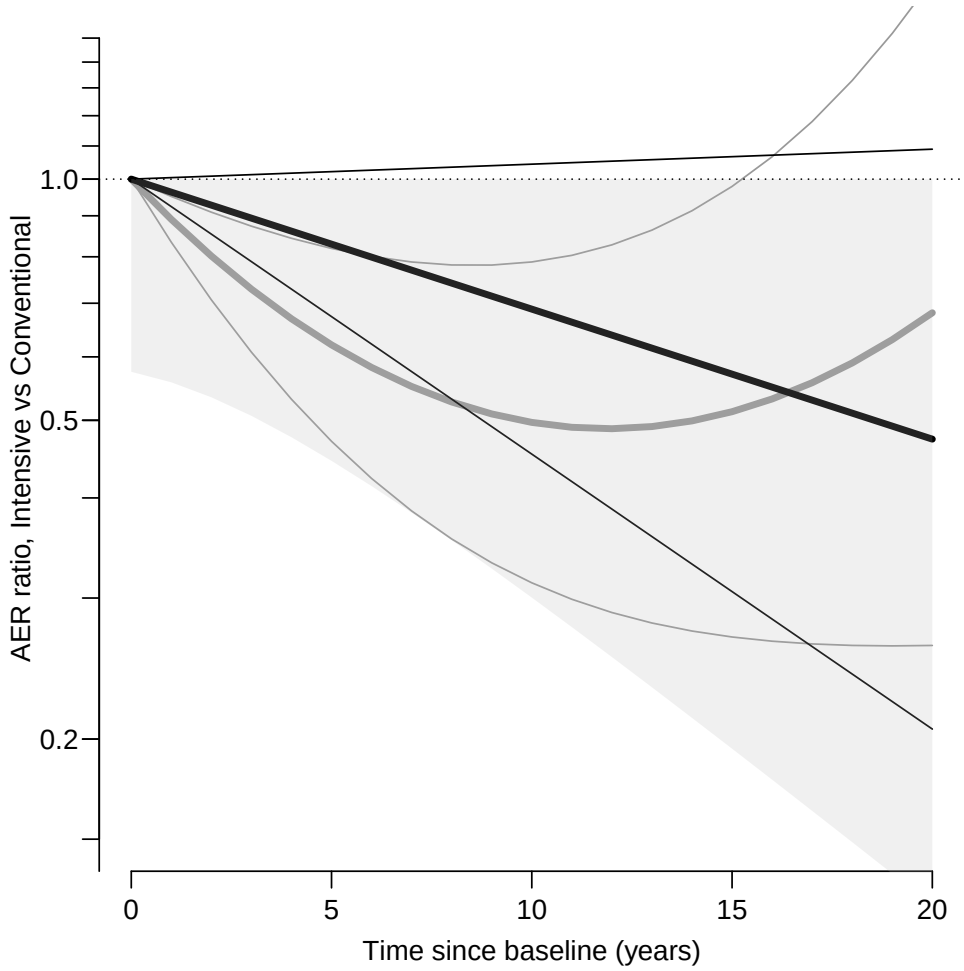

Figure 5.21: *Estimated mean AER-ratio between intensive and conventional by time since baseline by two different models; one with a linear trend and one with a quadratic trend, both assuming identical mean AER levels at baseline (black lines). Thick lines are the mean AER ratio; thin lines the 95% confidence intervals. The gray area represents the between-person-variation (from the quadratic model): The height of the area is the median ratio of AER between two randomly selected persons from the same allocation group.*

```
> library( Epi )
> clear()
> load( file="../data/mic.Rda" )
> source( "usefuns.R" )
```

We generate tables of states (`aer_level`) and dates of ascertainment of these, and generates a Lexis object with imputed follow-up times:

```
> levels( mic$aer_level )
[1] "Normoalbuminuria" "Microalbuminuria" "Macroalbuminuria"
> rt <- with( mic, tapply( as.integer(aer_level), list(ID,visit), mean ) )
> rd <- with( mic, tapply( doVis, list(ID,visit), mean ) )
> set.seed( 699817648 )
> alb <- make.dlex( rt, rd, incr=FALSE )
```

Since everyone is micro-albuminuric at start, we could either restrict transitions to be only increasing (only increase to macro), or we could allow transitions both ways. We have

chosen the latter to capture the possible beneficial effect of the intensive treatment. However we are formally making a mistake by assuming that transitions only occur in *one* direction between visits, which may not always be the case in reality.

```
> levels( alb )
[1] "1"       "2"       "3"       "Dead(1)" "Dead(2)" "Dead(3)"
> levels( alb$lex.Cst ) <-
+ levels( alb$lex.Xst ) <- c("Normo","Micro","Macro","D(Norm)","D(Mic)","D(Mac)")
> summary( alb )
Transitions:
     To
From    Normo Micro Macro D(Norm) D(Mic) D(Mac)  Records:  Events: Risk time:  Persons:
  Normo   154    36     0      21      0      0       211       57     572.90        69
  Micro    75   336    70       0     40      0       521      185    1419.29       160
  Macro     0    23   103       0      0     32       158       55     425.66        64
  Sum     229   395   173      21     40     32       890      297    2417.85       160
```

Here we plot the states, and the transitions between them:

```
> clr <- c("forestgreen",heat.colors(5)[c(3,1)])
> clr <- c(clr,rgb(t(col2rgb(clr)*0.6+255*0.4),max=255))
> parr <- rep(0.6,7)
> parr[c(2,5,7)] <- 0.25
> par( mfrow=c(1,2), oma=c(0,0,0,0) )
> for( lv in levels(alb$allocation) )
+    {
+ boxes( subset(alb,allocation==lv),
+        boxpos=list(x=rep(c(20,80),each=3)+6-12*(lv=="Intensive"),
+                    y=rep(seq(85,5,,3),2)),
+        show.BE="noz", scale.R=100,
+        col.bg=clr, pos.arr=parr,
+        col.border=c(clr[1:3],rep("black",3)),
+        col.txt=rep(c("white","black"),each=3) )
+ text(10+5-10*(lv=="Intensive"),98,lv,cex=1.5,font=2,adj=0)
+    }
```

The analysis of rates can be performed by a single Cox model simultaneously for all transitions. However, since we here (as opposed to what is the case for the retinopathy and neuropathy data) have multiple possible *exits* from some states, the analysis must be performed on the *stacked* version of the dataset:

```
> stalb <- stack( alb )
> str( stalb )
Classes 'stacked.Lexis' and 'data.frame':        2301 obs. of  16 variables:
 $ lex.id    : num  1 1 1 1 2 2 6 6 6 7 ...
 $ per       : num  1997 1997 2006 2007 1995 ...
 $ lex.dur   : num  0.435 1.497 0.575 0.175 0.45 ...
 $ lex.Cst   : Factor w/ 6 levels "Normo","Micro",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ lex.Xst   : Factor w/ 6 levels "Normo","Micro",..: 1 2 1 2 1 2 1 1 2 1 ...
 $ lex.Tr    : Factor w/ 7 levels "Normo->Micro",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ lex.Fail  : logi  FALSE TRUE FALSE TRUE FALSE TRUE ...
 $ sex       : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 2 2 2 1 ...
 $ doBth     : num  1932 1932 1932 1932 1947 ...
 $ doVis     : num  1993 1993 1993 1993 1993 ...
 $ doBlind   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ doEnd     : num  2015 2015 2015 2015 2015 ...
 $ doDth     : num  NA NA NA NA NA ...
 $ allocation: Factor w/ 2 levels "Intensive","Conventional": 1 1 1 1 1 1 2 2 2 2 ...
 $ age       : num  64.4 64.8 73.8 74.4 48.1 ...
 $ tfe       : num  3.34 3.78 12.78 13.36 1.57 ...
 - attr(*, "time.scales")= chr  "per" "age" "tfe"
> with( stalb, ftable(table(lex.Tr,allocation,lex.Fail),col.vars=2:3) )
```

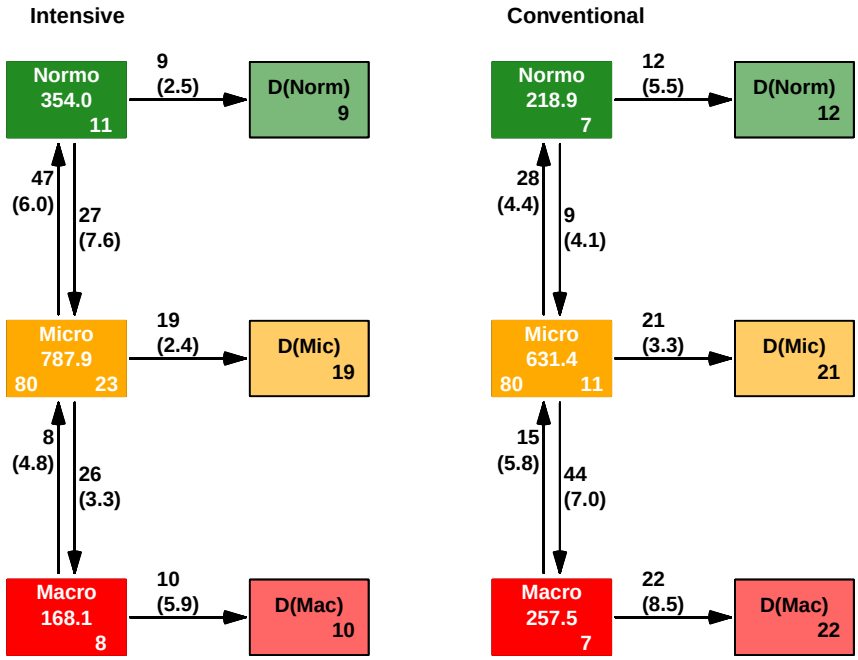

Figure 5.22: *Transitions between states in the two treatment groups, allowing transitions to take place in both directions, but only in one direction between visits. The number in the center of the boxes is the person-years, the numbers at the bottom is the number of patients starting, resp. ending their follow-up in each box, and the numbers on the arrows are the number of transitions and the overall transition rates per 100 PY.*

```
              allocation Intensive      Conventional
              lex.Fail       FALSE TRUE        FALSE TRUE
lex.Tr
Normo->Micro                   103   27           72    9
Normo->D(Norm)                 121    9           69   12
Micro->Normo                   227   47          219   28
Micro->Macro                   248   26          203   44
Micro->D(Mic)                  255   19          226   21
Macro->Micro                    51    8           84   15
Macro->D(Mac)                   49   10           77   22
```

We then fit the Cox model, using a separate baseline as well as a separate intervention effect for each of the transitions. Note that we are using the time since entry as the timescale, and not correcting for age or sex.

```
> m1 <- coxph( Surv(tfe,tfe+lex.dur,lex.Fail) ~ strata(lex.Tr) + allocation:lex.Tr,
+              data = stalb )
> HR <- cbind( ci.exp( m1, pval=TRUE ),
+            1/ci.exp( m1 ) )[c(6,3,1,4,2,5,7),c(1:3,5,7,6,4)]
> colnames(HR)[c(1,4)] <- c("Int vs Conv","Conv vs Int")
> rownames(HR) <- gsub("allocationIntensive:lex.Tr","",rownames(HR))
> round( HR, 3 )
              Int vs Conv  2.5% 97.5% Conv vs Int 97.5%  2.5%     P
Macro->Micro        0.967 0.407 2.297       1.034 0.435 2.457 0.939
Micro->Normo        1.612 1.007 2.579       0.620 0.388 0.993 0.047
Normo->Micro        1.806 0.848 3.843       0.554 0.260 1.179 0.125
Micro->Macro        0.514 0.315 0.840       1.944 1.190 3.175 0.008
Normo->D(Norm)      0.455 0.192 1.082       2.197 0.925 5.219 0.075
Micro->D(Mic)       0.640 0.341 1.202       1.562 0.832 2.933 0.166
Macro->D(Mac)       0.670 0.316 1.422       1.493 0.703 3.167 0.297
```

Here the first 3 columns refer to the intensive vs conventional, and the the last three to the opposite comparison. We see that the only formally significant effect of intervention is the halving of the transition from micro- to macro-albuminuria, HR 0.49 (0.30;0.79), p=0.004 — the other three transition rates do not differ significantly between groups, although there is a tendency that the intervention group also have higher transition rates between micro- and normo- albuminuria. The mortality rates are all roughly halved in the intensive group.

```
> Cests <- ci.exp( m1 )[c(6,3,1,4,2,5,7),c(1,3,2)]
> rownames( Cests ) <- gsub("allocationIntensive:lex.Tr","",rownames(Cests))
> plotEst( Cests, y=c(8:7+0.5,6:5,3:1), xlog=T, vref=1, grid=c(1:9/10,1:5),
+          xlab="Intensive vs. Conventional HR" )
```

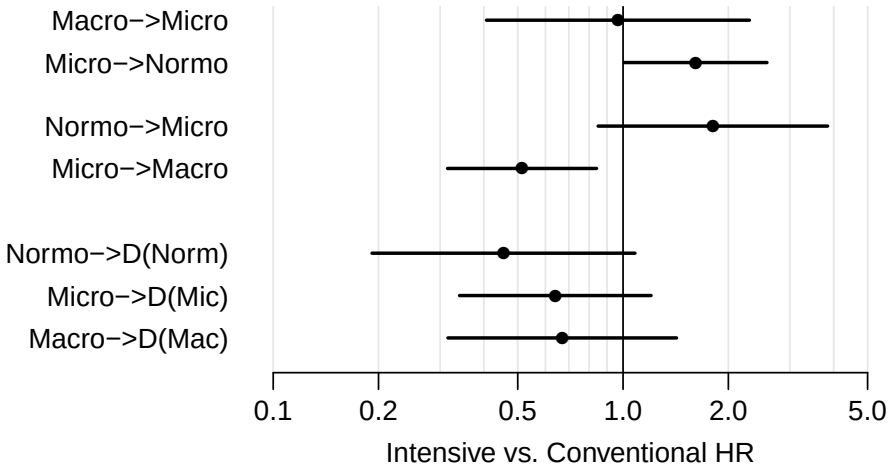

Figure 5.23: *Hazard ratios between Intensive and Conventional for the 7 transitions in figure 5.22, based on separate Cox models with time since baseline as time-scale and intervention as the only covariate. The models thus assume proportional hazards between the two intervention groups for each transition.*

However these numbers are just from a single imputation of the transition times, so should be taken with some caution.

5.4.3 Sensitivity analysis

In order to get a handle on the uncertainty introduced by the imputation of transition times we wrap up the entire imputation and analysis in a function that produces the estimates and their variance-covariance as result. Collecting these we can produce a summary estimate with corrected standard errors and moreover give an estimate of how large the contribution of the imputation variance is relative to the model.

Note that we only take the diagonal of the variance matrix, since it by the structure of the model *is* diagonal — the model fitting is tantamount to fitting each of the 7 transitions separately.

```
> alb1 <-
+ function()
+     {
+ alb <- make.dlex( rt, rd, incr=FALSE )
```

```
+ levels( alb$lex.Cst ) <-
+ levels( alb$lex.Xst ) <- c("Normo","Micro","Macro","D(Norm)","D(Mic)","D(Mac)")
+ stalb <- stack( alb )
+ m1 <- coxph( Surv(tfe,tfe+lex.dur,lex.Fail) ~ strata(lex.Tr) + allocation:lex.Tr,
+              data = stalb )
+ zz <- ci.lin( m1 )[,1:2]
+ zz
+     }
> system.time( res <- alb1() )
   user  system elapsed
  0.439   0.004   0.444
```

With this in place we set up an array to hold the results of analyses of data from multiple different imputations:

```
> nsim <- 400
> impalb <- NArray( c( list(imp=1:nsim), dimnames(res) ) )
> dimnames(impalb)[[2]] <- gsub("allocationConventional:lex.Tr","",dimnames(impalb)[[2]])
> str( impalb )
 logi [1:400, 1:7, 1:2] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 3
  ..$ imp: chr [1:400] "1" "2" "3" "4" ...
  ..$    : chr [1:7] "allocationIntensive:lex.TrNormo->Micro" "allocationIntensive:lex.TrNormo->D(No
  ..$    : chr [1:2] "Estimate" "StdErr"
> system.time( for(i in 1:nsim) impalb[i,,] <- alb1() )
   user  system elapsed
193.046   0.059 193.480
> vv <-  cbind( apply( impalb[,,2]  , 2, sd ),
+               apply( impalb[,,1]  , 2, var ),
+               apply( impalb[,,2]^2, 2, mean ) )
> colnames( vv ) <- c("SD(sd)","VAR(est)","MEAN(var)")
> round( cbind(vv,"VAR(est)/MEAN(var)"=vv[,1]/vv[,3]), 4 )
                                        SD(sd) VAR(est) MEAN(var) VAR(est)/MEAN(var)
allocationIntensive:lex.TrNormo->Micro  0.0003   0.0026    0.1487             0.0022
allocationIntensive:lex.TrNormo->D(Norm) 0.0004   0.0027    0.1953             0.0022
allocationIntensive:lex.TrMicro->Normo  0.0002   0.0004    0.0574             0.0034
allocationIntensive:lex.TrMicro->Macro  0.0005   0.0009    0.0624             0.0075
allocationIntensive:lex.TrMicro->D(Mic) 0.0009   0.0013    0.1031             0.0090
allocationIntensive:lex.TrMacro->Micro  0.0016   0.0061    0.1948             0.0083
allocationIntensive:lex.TrMacro->D(Mac) 0.0012   0.0076    0.1477             0.0083
```

Thus for all practical purposes it does not matter whether we use the average over multiple imputations or just a single imputation; the variance will be inflated by about 3% at most, and thus the s.e.s with $\sqrt{1.03} - 1 = 1.015 - 1 = 1.5\%$ at most. Moreover, the estimate bases on a single imputation of time will with 99% probability be off by less than 5% of the standard error.

Thus we can safely rely on the results derived from the initial imputation, and we shall therefore just use this initially imputed transition times in further analyses.

5.4.4 Prediction of state occupancy

For the sake of illustration we would like to show the state occupancy, and in particular we would like to show how many die from different states — basically we want to see how many persons are in each of the states illustrated in figure 5.22 as a function of time since baseline.

If we want to illustrate how patients fare between the states we will need models for each of the 7 transitions; here we shall fit Poisson-models smoothing the trends and taking age and sex into account. The total number of events for the transitions vary from 20 to 70.

First we split the follow-up in 1-month intervals:

```
> summary( alb )
Transitions:
     To
From    Normo Micro Macro D(Norm) D(Mic) D(Mac)  Records:  Events: Risk time:  Persons:
  Normo   154    36     0      21      0      0       211       57     572.90        69
  Micro    75   336    70       0     40      0       521      185    1419.29       160
  Macro     0    23   103       0      0     32       158       55     425.66        64
  Sum     229   395   173      21     40     32       890      297    2417.85       160
> timeScales( alb )
[1] "per" "age" "tfe"
> spa <- splitLexis( alb, seq(0,24,1/12), "tfe" )
> summary( spa )
Transitions:
     To
From    Normo Micro Macro D(Norm) D(Mic) D(Mac)  Records:  Events: Risk time:  Persons:
  Normo  7024    36     0      21      0      0      7081       57     572.90        69
  Micro    75 17289    70       0     40      0     17474      185    1419.29       160
  Macro     0    23  5211       0      0     32      5266       55     425.66        64
  Sum    7099 17348  5281      21     40     32     29821      297    2417.85       160
```

Then we can fit models for each of the transitions, first we determine the position of the knots for the underlying splines:

```
> ( d.kn <- with( subset( spa, substr(lex.Xst,1,1)=="D" ),
+                 c(0,quantile( tfe+lex.dur, probs=0:2/3)) ) )
                  0% 33.33333% 66.66667%
 0.000000  1.637235  8.276523 13.013005
> tkn <- addmargins( as.table( rbind(
+ ( mino.kn <- with( subset( spa, lex.Cst=="Micro" & lex.Xst=="Normo" ),
+                 c(0,quantile( tfe+lex.dur, probs=0:2/3)) ) ),
+ ( mima.kn <- with( subset( spa, lex.Cst=="Micro" & lex.Xst=="Macro" ),
+                 c(0,quantile( tfe+lex.dur, probs=0:2/3)) ) ),
+ ( nomi.kn <- with( subset( spa, lex.Cst=="Normo" & lex.Xst=="Micro" ),
+                 c(0,quantile( tfe+lex.dur, probs=0:2/3)) ) ),
+ ( mami.kn <- with( subset( spa, lex.Cst=="Macro" & lex.Xst=="Micro" ),
+                 c(0,quantile( tfe+lex.dur, probs=0:2/3)) ) ) ) ),
+                    margin=1, FUN=mean )
> round( tkn, 2 )
             0% 33.33333% 66.66667%
A      0.00  0.01      1.70      4.60
B      0.00  0.18      2.79      6.82
C      0.00  2.18      6.39     11.28
D      0.00  1.68      4.55      8.65
mean   0.00  1.01      3.86      7.84
```

There is some variation of the quantiles of the events between the different transitions, but we use the same set of knots anyway — the mean:

```
> ( m.kn <- tkn["mean",] )
                  0% 33.33333% 66.66667%
 0.000000  1.011906  3.858181  7.839780
```

We can now fit the 5 different models; 4 for the transitions and one for the mortality rates. For the mortality rates we first check if the rates are actually proportional.

```
> ( dth <- levels(spa)[4:6] )
[1] "D(Norm)" "D(Mic)"  "D(Mac)"
> m0 <- glm( ( lex.Xst %in% dth ) ~
+            Ns( tfe, knots=d.kn) + allocation + lex.Cst + age + sex,
+            offset = log(lex.dur),
+            family=poisson,
+            data = spa )
> mi  <- update( m0, . ~ . + Ns( tfe, knots=d.kn):lex.Cst )
> mI  <- update( m0, . ~ . + Ns( tfe, knots=d.kn):allocation )
> miI <- update( mi, . ~ . + Ns( tfe, knots=d.kn):allocation )
> round( ci.exp( m0 ), 3 )
```

```
                      exp(Est.)  2.5%    97.5%
(Intercept)               0.000 0.000    0.000
Ns(tfe, knots = d.kn)1    6.719 1.361   33.173
Ns(tfe, knots = d.kn)2   43.131 0.408 4561.508
Ns(tfe, knots = d.kn)3    2.132 0.854    5.323
allocationConventional    1.800 1.181    2.744
lex.CstMicro              0.808 0.475    1.377
lex.CstMacro              1.341 0.763    2.355
age                       1.107 1.070    1.145
sexMale                   1.392 0.840    2.308
> round( ci.exp( mI ), 3 )
                                                exp(Est.)  2.5%      97.5%
(Intercept)                                         0.000 0.000      0.000
Ns(tfe, knots = d.kn)1                              2.271 0.309     16.668
Ns(tfe, knots = d.kn)2                             13.786 0.040   4792.915
Ns(tfe, knots = d.kn)3                              0.945 0.286      3.122
allocationConventional                              0.309 0.002     38.358
lex.CstMicro                                        0.807 0.472      1.381
lex.CstMacro                                        1.343 0.762      2.367
age                                                 1.108 1.071      1.146
sexMale                                             1.398 0.844      2.316
Ns(tfe, knots = d.kn)1:allocationConventional       8.366 0.296    236.626
Ns(tfe, knots = d.kn)2:allocationConventional      16.327 0.001 278783.818
Ns(tfe, knots = d.kn)3:allocationConventional       4.484 0.734     27.410
> AoV <- anova( miI, mi, m0, mI, miI, test="Chisq" )[-1,-(1:2)]
> rownames( AoV ) <- c("t*allo|t*st","t*st","t*allo","t*st|t*allo")
> round( AoV, 3 )
            Df Deviance Pr(>Chi)
t*allo|t*st -3   -2.551    0.466
t*st        -6  -14.169    0.028
t*allo       3    3.119    0.374
t*st|t*allo  6   13.601    0.034
```

Thus it is pretty clear that we cannot assume proportional hazards of death between different levels of albuminuria, but equally clear that mortality is proportional between treatment groups, thus the model `mi` is the model to use. We will incorporate this in the mortality model, so we now only need the 4 models linking the different levels of albuminuria;

```
> mino <- glm( ( lex.Xst %in% "Normo" ) ~
+              Ns( tfe, knots=m.kn) + allocation + age + sex,
+              offset = log(lex.dur),
+              family=poisson,
+              data = subset(spa,lex.Cst=="Micro" ) )
> nomi <- update( mino,  ( lex.Xst %in% "Micro" ) ~ . , data = subset(spa,lex.Cst=="Normo" ) )
> mima <- update( mino,  ( lex.Xst %in% "Macro" ) ~ . , data = subset(spa,lex.Cst=="Micro" ) )
> mami <- update( mino,  ( lex.Xst %in% "Micro" ) ~ . , data = subset(spa,lex.Cst=="Macro" ) )
```

The models we fitted are proportional hazards models, so we briefly make formal tests for proportional hazards along the timescale:

```
> ph3 <- rbind(
+ anova( mino, update(mino,.~.+Ns(tfe,knots=m.kn):allocation), test="Chisq" )[2,-(1:2)],
+ anova( nomi, update(nomi,.~.+Ns(tfe,knots=m.kn):allocation,maxit=100), test="Chisq" )[2,-(1:2)],
+ anova( mima, update(mima,.~.+Ns(tfe,knots=m.kn):allocation), test="Chisq" )[2,-(1:2)],
+ anova( mami, update(mami,.~.+Ns(tfe,knots=m.kn):allocation), test="Chisq" )[2,-(1:2)])
> ph1 <- rbind(
+ anova( mino, update(mino,.~.+tfe:allocation), test="Chisq" )[2,-(1:2)],
+ anova( nomi, update(nomi,.~.+tfe:allocation,maxit=100), test="Chisq" )[2,-(1:2)],
+ anova( mima, update(mima,.~.+tfe:allocation), test="Chisq" )[2,-(1:2)],
+ anova( mami, update(mami,.~.+tfe:allocation), test="Chisq" )[2,-(1:2)])
> rownames( ph3 ) <- rownames( ph1 ) <- c("mino","nomi","mima","mami")
> round( cbind(ph3,ph1), 3 )
```

```
     Df Deviance Pr(>Chi) Df Deviance Pr(>Chi)
mino  3    0.748    0.862  1    0.661    0.416
nomi  3    4.273    0.233  1    0.000    0.986
mima  3    5.758    0.124  1    0.864    0.353
mami  3    8.269    0.041  1    1.630    0.202
```

Thus we see that there is no compelling indication of non-proportional hazards for any of the transitions, neither with a 3 d.f. (except possibly for the transition macro to micro) nor a 1 d.f. alternative.

We then update the HR collector for peripheral neuropathy

```
> mima0 <- update( mima, . ~ . - age - sex )
> zz <- rbind( ci.exp( mima0, subset="allo", pval=TRUE ),
+              ci.exp( mima , subset="allo", pval=TRUE ) )
> zz[,1:3] <- 1/zz[,c(1,3,2)]
> load( file="../data/mainCI.Rda" )
> dimnames( mainCI )[[1]][10]
[1] "Macroalbuminuria"

> mainCI[10,,] <- zz
> round( ftable(mainCI,row.vars=1), 3 )
                      model      raw                      age/sex
                      what  Estimate  2.5% 97.5%     P Estimate  2.5% 97.5%     P
outcome
All cause mortality             0.551 0.364 0.835 0.005    0.526 0.346 0.800 0.003
CVD mortality                   0.379 0.191 0.754 0.006    0.353 0.177 0.705 0.003
non-CVD mortality               0.703 0.413 1.197 0.195    0.686 0.401 1.172 0.168
Death or 1st CVD                   NA    NA    NA    NA       NA    NA    NA    NA
Death | CVD state               0.835 0.536 1.301 0.425    0.874 0.564 1.355 0.548
CVD event | CVD state           0.552 0.394 0.772 0.001    0.575 0.412 0.803 0.001
Retinopthy progression          0.668 0.507 0.881 0.004    0.673 0.511 0.887 0.005
Autonomic neuropathy            0.594 0.397 0.890 0.011    0.573 0.382 0.859 0.007
Peripheral neuropathy           1.120 0.707 1.774 0.630    1.101 0.694 1.747 0.683
Macroalbuminuria                0.516 0.316 0.842 0.008    0.495 0.302 0.811 0.005

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

We can then derive the HRs associated with treatment in the different models:

```
> Pests <- Cests*0
> Pests["Macro->Micro"  ,] <- 1/ci.exp(mami,subset="allo")
> Pests["Micro->Normo"  ,] <- 1/ci.exp(mino,subset="allo")
> Pests["Normo->Micro"  ,] <- 1/ci.exp(nomi,subset="allo")
> Pests["Micro->Macro"  ,] <- 1/ci.exp(mima,subset="allo")
> Pests["Normo->D(Norm)",] <- 1/ci.exp(mi,subset="allo")
> Pests["Micro->D(Mic)" ,] <- 1/ci.exp(mi,subset="allo")
> Pests["Macro->D(Mac)" ,] <- 1/ci.exp(mi,subset="allo")
```

and we can then plot these comparing with the raw ones from the Cox model

```
> par( mfrow=c(1,1), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6 )
> plotEst( Pests, y=c(8:7+0.5,6:5,3:1)+0.1, xlog=T, vref=1, grid=c(1:9/10,1:5),
+          xlab="Intensive vs. Conventional HR" )
> linesEst( Cests, y=c(8:7+0.5,6:5,3:1)-0.1, xlog=T, col=gray(0.6) )
```

5.4.4.1 State occupancy

With models for all transitions in place we can now set up a transition object to be used in simulating from an initial data frame:

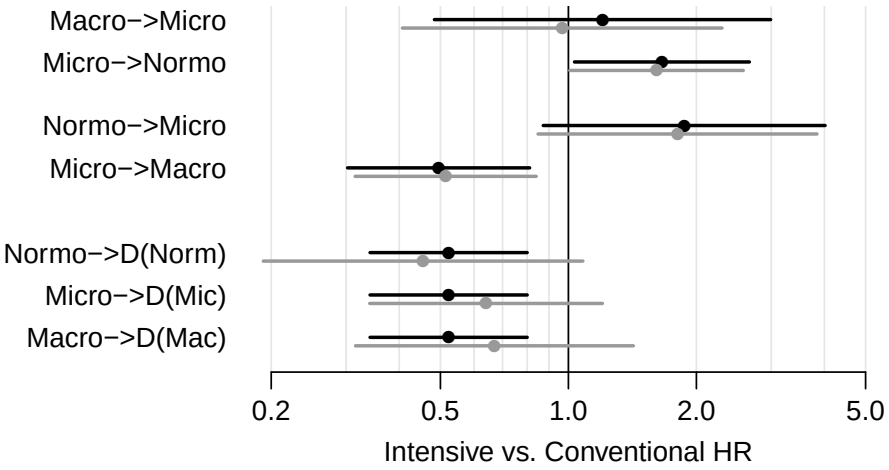

Figure 5.24: *Hazard ratios between Intensive and Conventional for the 7 transitions in figure 5.22, based on Poisson models for transitions (black) based on separate models for transition between albuminuria states and a single model for the mortalities. Models use time since baseline as time-scale effects of intervention and intervention, age and sex as covariates. The gray bars are the estimates from the simple Cox-models.*

```
> TR <- list( Normo = list( Micro = nomi,
+                           "D(Norm)" = mi),
+             Micro = list( Normo = mino,
+                           Macro = mima,
+                           "D(Mic)" = mi),
+             Macro = list( Micro = mami,
+                           "D(Mac)" = mi) )
> lapply( TR, names )
$Normo
[1] "Micro"   "D(Norm)"

$Micro
[1] "Normo"  "Macro"  "D(Mic)"

$Macro
[1] "Micro"  "D(Mac)"
> ini <- subset( alb, tfe==0,
+                select=c(timeScales(alb),"lex.id","lex.Cst","allocation","sex") )
> ini <- rbind( cbind( ini, allocation="Intensive" ),
+               cbind( ini, allocation="Conventional" ) )
> ini$lex.Cst <- factor( ini$lex.Cst, levels=levels( ini$lex.Cst )[1:3] )
> levels( ini )
[1] "Normo" "Micro" "Macro"
> str( ini )
Classes 'Lexis' and 'data.frame':        320 obs. of  7 variables:
 $ per       : num  1993 1993 1993 1993 1993 ...
 $ age       : num  61.1 46.5 50 48.6 57.2 ...
 $ tfe       : num  0 0 0 0 0 0 0 0 0 0 ...
 $ lex.id    : num  1 2 3 4 5 6 7 8 9 10 ...
 $ lex.Cst   : Factor w/ 3 levels "Normo","Micro",..: 2 2 2 2 2 2 2 2 2 2 ...
 $ allocation: Factor w/ 2 levels "Intensive","Conventional": 1 1 2 2 2 2 2 1 1 1 ...
 $ sex       : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 1 2 2 2 ...
 - attr(*, "breaks")=List of 3
  ..$ per: NULL
  ..$ age: NULL
  ..$ tfe: NULL
 - attr(*, "time.scales")= chr  "per" "age" "tfe"
 - attr(*, "time.since")= chr  "" "" ""
```

We can now simulate the follow-up through the states for a cohort 100 times bigger then the Steno2 study population:

```
> system.time( Salb <- simLexis( TR, ini, N=100, time.pts=0:100/4 ) )
   user  system elapsed
130.483  11.415 142.130
> summary( Salb, by="allocation" )
$Intensive

Transitions:
     To
From    Normo Micro Macro D(Mic) D(Mac) D(Norm)  Records:  Events: Risk time:  Persons:
  Normo  1774  5916     0      0      0    2034      9724     7950   79368.76      8555
  Micro  9724  4150  5676   4189      0       0     23739    19589  175520.39     16000
  Macro     0  1823  1224      0   2629       0      5676     4452   39385.69      5240
  Sum   11498 11889  6900   4189   2629    2034     39139    31991  294274.84     16000

$Conventional

Transitions:
     To
From    Normo Micro Macro D(Mic) D(Mac) D(Norm)  Records:  Events: Risk time:  Persons:
  Normo  1252  2050     0      0      0    2357      5659     4407   53769.82      5428
  Micro  5659  1858  8689   4701      0       0     20907    19049  134062.58     16000
  Macro     0  2857   961      0   4871       0      8689     7728   53513.07      7690
  Sum    6911  6765  9650   4701   4871    2357     35255    31184  241345.47     16000
> save( Salb, file="../data/Salb.Rda" )


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

We then compute the number of persons in each state at 3-month intervals

```
> pm <- c(1:3,5,4,6)
> levels(Salb)[pm]
[1] "Normo"   "Micro"   "Macro"   "D(Mac)"  "D(Mic)"  "D(Norm)"
> Pi <- pState( nState( subset(Salb,allocation=="Intensive"),
+                      at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=pm )
> Pc <- pState( nState( subset(Salb,allocation=="Conventional"),
+                      at=seq(0,20,1/4), from=0, time.scale="tfe" ), perm=pm )
> head( Pi )
      State
when       Normo     Micro     Macro    D(Mac) D(Mic) D(Norm)
  0    0.0000000 1.0000000 1.0000000 1.0000000      1       1
  0.25 0.0275625 0.9916875 0.9998750 0.9998750      1       1
  0.5  0.0577500 0.9842500 0.9992500 0.9992500      1       1
  0.75 0.0903125 0.9768750 0.9986250 0.9986250      1       1
  1    0.1208125 0.9676250 0.9980000 0.9980000      1       1
  1.25 0.1513125 0.9589375 0.9974375 0.9974375      1       1


> albst <- function(rv=FALSE,leg=15){
+ par( mfrow=c(1,2), mar=c(0,3,0,0), oma=c(3,0,3,3-2*rv), las=1 )
+ plot( Pi, col=clr[c(1:3,6:4)], xlim=c(0,20), xlab="", yaxt="n",
+       col.lab="transparent" )
+ lines( as.numeric(rownames(Pi)), Pi[,"Macro"], lwd=4 )
+ if(is.numeric(leg)) text( rep(leg,6),
+                           Pi[paste(leg),]-diff(c(0,Pi[paste(leg),]))/2,
+                           colnames(Pi),
+                           col=rep(gray(1:0),each=3) )
+ axis( side=4, at=0:5/5 )
+ axis( side=4, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4, at=0:100  /100, tcl=-0.2, labels=NA )
```

```
+ mtext( "Probability", side=2, line=1, las=0 )
+ mtext( "Intensive", side=3, line=1 )
+
+ plot( Pc, col=clr[c(1:3,6:4)], xlim=if(rv) c(20,0) else c(0,20),
+       xlab="", yaxt="n", col.lab="transparent" )
+ lines( as.numeric(rownames(Pc)), Pc[,"Macro"], lwd=4 )
+ axis( side=4-2*rv, at=1:5/5, if(rv) labels=NA )
+ axis( side=4-2*rv, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4-2*rv, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4-2*rv, at=0:100  /100, tcl=-0.2, labels=NA )
+ mtext( "Conventional", side=3, line=1 )
+ mtext( "Time since baseline (years)", side=1, line=2, outer=TRUE )
+ }
> albst(rv=FALSE,leg=18)


> albst(rv=TRUE,leg=18)
```

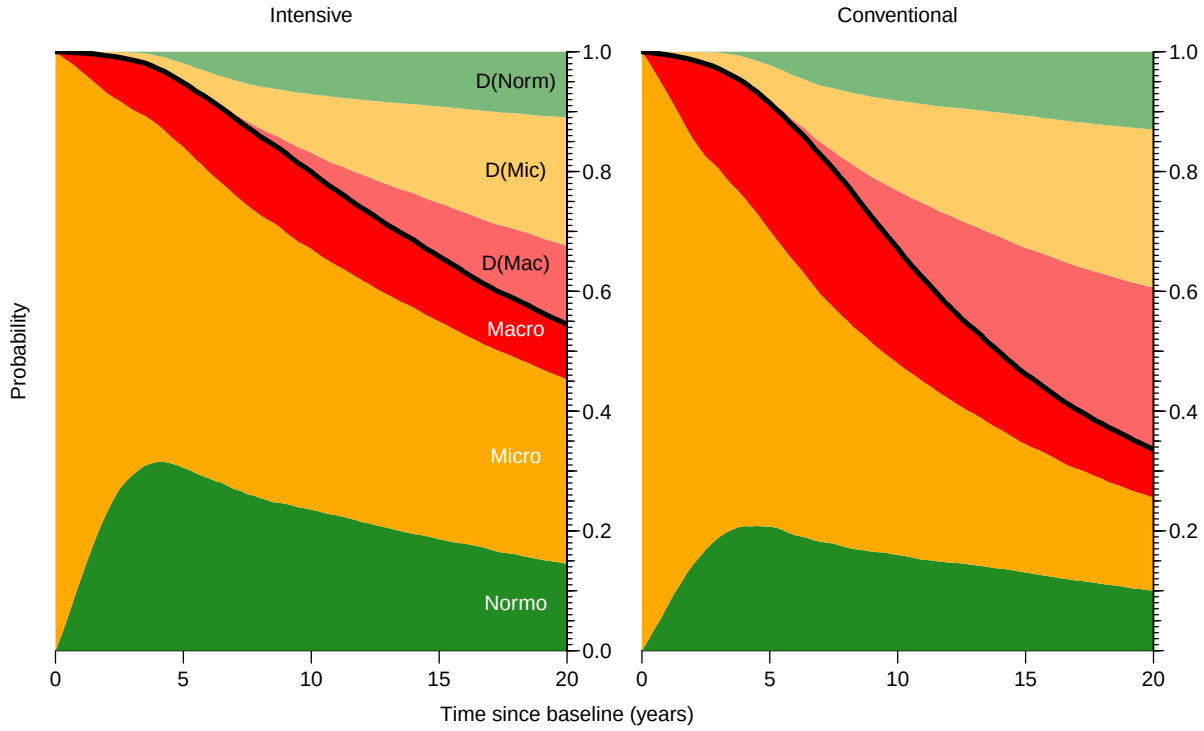

Figure 5.25: *Probability of being in each state (from bottom: Normo-, Micro- and Macro-albuminuria) as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*

The figures 5.25 and 5.26 show the predicted distribution of patients in Normo-, Micro- and Macro-albuminuria using assumptions that transitions occur at a simulated time between visits where status were assessed. Since we are modeling mortality conditional on albuminuria state, we see survival curves tat are sightly different from the ones seen in the overall mortality analysis.

Finally we save the state probability results:

```
> save( Pi, Pc, file="../data/pAlb.Rda" )
```

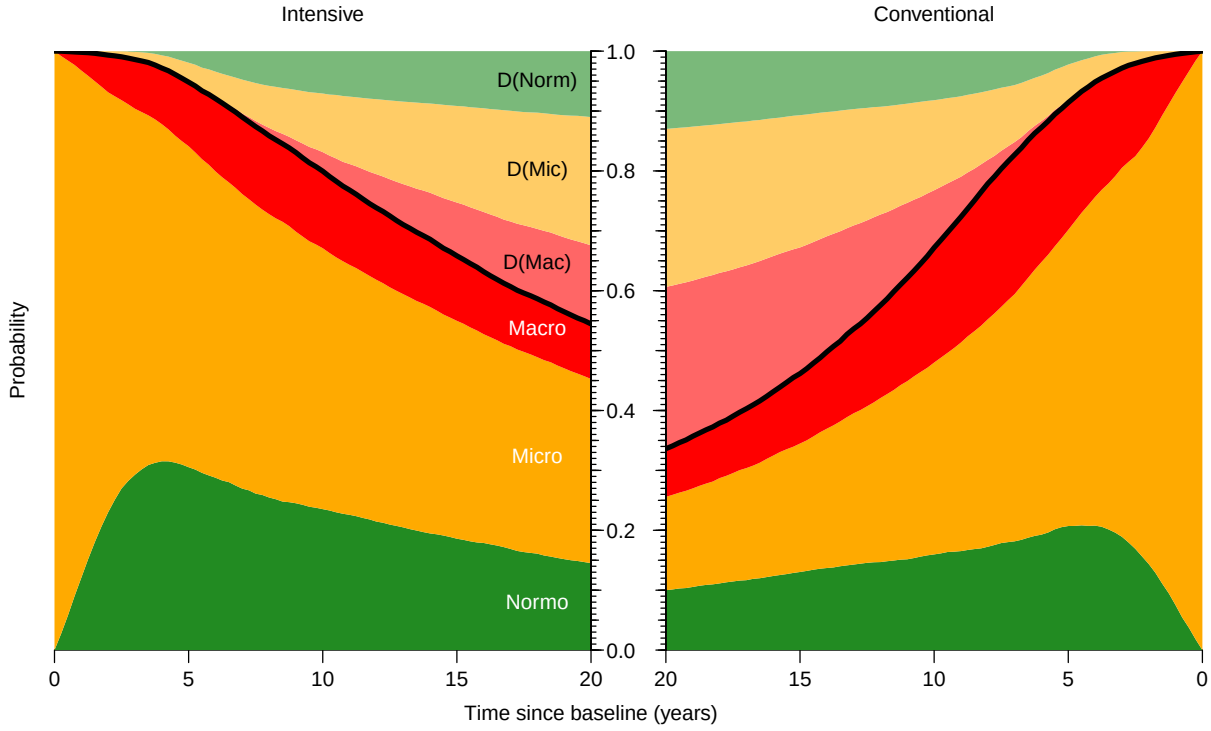

Figure 5.26: *Probability of being in each state (from bottom: Normo-, Micro- and Macro-albuminuria) as a function of time since baseline. The black curve is the estimated survival curve, and the pale colors in the dead states represent the states* from *which person have died.*
This is the same figure as 5.25, except that the x-axis of the right-hand panel is reversed.

5.5 Comparative plotting of microvascular complications

The relevant `pState` objects are:

- Autonomic neuropathy: `pAi`, `pAc`
- Peripheral neuropathy: `pPi`, `pPc`
- Retinopathy: `Ei`, `Ec` — the obejcts from models that only allows progression of retinopathy, using the empirical baseline distribution as starting point.

  The objects `rEi`, `rEc` are from the models that allow improvement of retinopathy, but they are likely to be shaky beacuse of the small number of imorovement transitions.
- Albumunuria: `Pi`, `Pc`

First we devise a function that plots state distribution from the two groups together

```
> mk.clr <-
+ function( clr )
+ {
+ clr <- c("forestgreen",clr)
+ c(clr,rgb(t(col2rgb(rev(clr))*0.6+255*0.4),max=255))
+ }
> micst <-
+ function( Pi, Pc, rv=TRUE, xl=c(0,20), clr="orange", xaxt="n" )
+ {
+ nc <- ncol(Pi)/2
+ clr <- mk.clr( clr )
+ plot( Pi, col=clr, xlim=xl, xlab="", yaxt="n",
+           col.lab="transparent", xaxt=xaxt )
+ lines( as.numeric(rownames(Pi)), Pi[,nc], lwd=4 )
+ Pi <- cbind(0,Pi)
+ axis( side=2, lend=1, lwd=0, lwd.ticks=1, at=1:10/10, labels=NA )
+ axis( side=2, lend=1, lwd=0, lwd.ticks=1, at=((1:5)*2-1)/10 )
+ axis( side=4, lend=1, lwd=0, lwd.ticks=1, at=((1:5)*2-1)/10 )
+ axis( side=4, lend=1, lwd=0, lwd.ticks=1, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4, lend=1, lwd=0, lwd.ticks=1, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4, lend=1, lwd=0, lwd.ticks=1, at=0:100  /100, tcl=-0.2, labels=NA )
+
+ plot( Pc, col=clr, xlim=if(rv) xl[2:1] else xl,
+       xlab="", yaxt="n", col.lab="transparent", xaxt=xaxt )
+ lines( as.numeric(rownames(Pc)), Pc[,nc], lwd=4 )
+ axis( side=2+2*rv, lend=1, lwd=0, lwd.ticks=1, at=1:10/10, labels=NA )
+ axis( side=2+2*rv, lend=1, lwd=0, lwd.ticks=1, at=((1:5)*2-1)/10 )
+ axis( side=4-2*rv, lend=1, lwd=0, lwd.ticks=1, at=((1:5)*2-1)/10, if(rv) labels=NA )
+ axis( side=4-2*rv, lend=1, lwd=0, lwd.ticks=1, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4-2*rv, lend=1, lwd=0, lwd.ticks=1, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4-2*rv, lend=1, lwd=0, lwd.ticks=1, at=0:100  /100, tcl=-0.2, labels=NA )
+ clr <<- clr
+ }
> legpl <-
+ function( pP, tit, lnam=colnames(pP) )
+   {
+ plot( NA, ylim=c(0,100), xlim=c(0,100), axes=FALSE, ylab="" )
+ nst <- ncol(pP)/2
+ stn <- lnam[1:nst]
+ text( 5, 90, tit, adj=0, font=2, cex=1.5 )
+ text( rep(5,nst), 85-(nst:1)*10, stn, font=2, cex=1.5, adj=0, col=clr[1:nst] )
+   }
```

With these functions in place we can plot cumulative risks in the two groups:

```
> layout( matrix(1:12,4,3,byrow=T), widths=c(2,2,1.3) )
> par( # mfrow=c(4,3),
+      mar=c(1,2.5,0,0), oma=c(3,2,2.5,1), las=1 )
> micst( pAi, pAc, xl=c(0,20), clr="orange" )
> legpl( pAi, "Autonomic neuropathy", lnam=c("Stable","Progression") )
> micst( pPi, pPc, xl=c(0,20), clr="orange" )
> legpl( pPi, "Peripheral neuropathy", lnam=c("Stable","Progression") )
> micst( rEi, rEc, xl=c(0,20), clr=heat.colors(7)[5:1] )
> legpl( rEi, "Retinopathy", lnam=c("None","Minimal","Moderate","pre-Proliferative","Proliferative",
> micst( Pi, Pc, xl=c(0,20), clr=heat.colors(5)[c(3,1)], xaxt="s" )
> legpl( Pi, "Abuminuria" )
> mtext( c("Intensive","Conventional"), at=c(1,4)/8, side=3, line=1, las=0, outer=TRUE )
> mtext( "Probability", side=2, line=0, las=0, outer=TRUE )
> mtext( "Time since baseline (years)", side=1, line=1.5, outer=TRUE, at=0.42 )
```

5.6 Comparative plotting of microvascular complications

The relevant `pState` objects are:

- Autonomic neuropathy: `pAi`, `pAc`
- Peripheral neuropathy: `pPi`, `pPc`
- Retinopathy: `Ei`, `Ec` — the obejcts from models that only allows progression of retinopathy, using the empirical baseline distribution as starting point.

  The objects `rEi`, `rEc` are from the models that allow improvement of retinopathy, but they are likely to be shaky beacuse of the small number of imorovement transitions.
- Albumunuria: `Pi`, `Pc`

First we devise a function that plots state distribution from the two groups together

```
> mk.clr <-
+ function( clr )
+ {
+ clr <- c("forestgreen",clr)
+ c(clr,rgb(t(col2rgb(rev(clr))*0.6+255*0.4),max=255))
+ }
> micst <-
+ function( Pi, Pc, rv=TRUE, xl=c(0,20), clr="orange", xaxt="n" )
+ {
+ nc <- ncol(Pi)/2
+ clr <- mk.clr( clr )
+ plot( Pi, col=clr, xlim=xl, xlab="", yaxt="n",
+           col.lab="transparent", xaxt=xaxt )
+ lines( as.numeric(rownames(Pi)), Pi[,nc], lwd=4 )
+ Pi <- cbind(0,Pi)
+ axis( side=2, lend=1, lwd=0, lwd.ticks=1, at=1:10/10, labels=NA )
+ axis( side=2, lend=1, lwd=0, lwd.ticks=1, at=((1:5)*2-1)/10 )
+ axis( side=4, lend=1, lwd=0, lwd.ticks=1, at=((1:5)*2-1)/10 )
+ axis( side=4, lend=1, lwd=0, lwd.ticks=1, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4, lend=1, lwd=0, lwd.ticks=1, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4, lend=1, lwd=0, lwd.ticks=1, at=0:100  /100, tcl=-0.2, labels=NA )
+
+ plot( Pc, col=clr, xlim=if(rv) xl[2:1] else xl,
+       xlab="", yaxt="n", col.lab="transparent", xaxt=xaxt )
```

```
+ lines( as.numeric(rownames(Pc)), Pc[,nc], lwd=4 )
+ axis( side=2+2*rv, lend=1, lwd=0, lwd.ticks=1, at=1:10/10, labels=NA )
+ axis( side=2+2*rv, lend=1, lwd=0, lwd.ticks=1, at=((1:5)*2-1)/10 )
+ axis( side=4-2*rv, lend=1, lwd=0, lwd.ticks=1, at=((1:5)*2-1)/10, if(rv) labels=NA )
+ axis( side=4-2*rv, lend=1, lwd=0, lwd.ticks=1, at=0:10*10/100, tcl=-0.4, labels=NA )
+ axis( side=4-2*rv, lend=1, lwd=0, lwd.ticks=1, at=0:20*5 /100, tcl=-0.3, labels=NA )
+ axis( side=4-2*rv, lend=1, lwd=0, lwd.ticks=1, at=0:100  /100, tcl=-0.2, labels=NA )
+ clr <<- clr
+ }
> legpl <-
+ function( pP, tit, lnam=colnames(pP) )
+   {
+ plot( NA, ylim=c(0,100), xlim=c(0,100), axes=FALSE, ylab="" )
+ nst <- ncol(pP)/2
+ stn <- lnam[1:nst]
+ text( 5, 90, tit, adj=0, font=2, cex=1.5 )
+ text( rep(5,nst), 85-(nst:1)*10, stn, font=2, cex=1.5, adj=0, col=clr[1:nst] )
+   }
```

With these functions in place we can plot cumulative risks in the two groups:

```
> layout( matrix(1:12,4,3,byrow=T), widths=c(2,2,1.3) )
> par( # mfrow=c(4,3),
+      mar=c(1,2.5,0,0), oma=c(3,2,2.5,1), las=1 )
> micst( pAi, pAc, xl=c(0,20), clr="orange" )
> legpl( pAi, "Autonomic neuropathy", lnam=c("Stable","Progression") )
> micst( pPi, pPc, xl=c(0,20), clr="orange" )
> legpl( pPi, "Peripheral neuropathy", lnam=c("Stable","Progression") )
> micst( rEi, rEc, xl=c(0,20), clr=heat.colors(7)[5:1] )
> legpl( rEi, "Retinopathy", lnam=c("None","Minimal","Moderate","pre-Proliferative","Proliferative",
> micst( Pi, Pc, xl=c(0,20), clr=heat.colors(5)[c(3,1)], xaxt="s" )
> legpl( Pi, "Abuminuria" )
> mtext( c("Intensive","Conventional"), at=c(1,4)/8, side=3, line=1, las=0, outer=TRUE )
> mtext( "Probability", side=2, line=0, las=0, outer=TRUE )
> mtext( "Time since baseline (years)", side=1, line=1.5, outer=TRUE, at=0.42 )
```

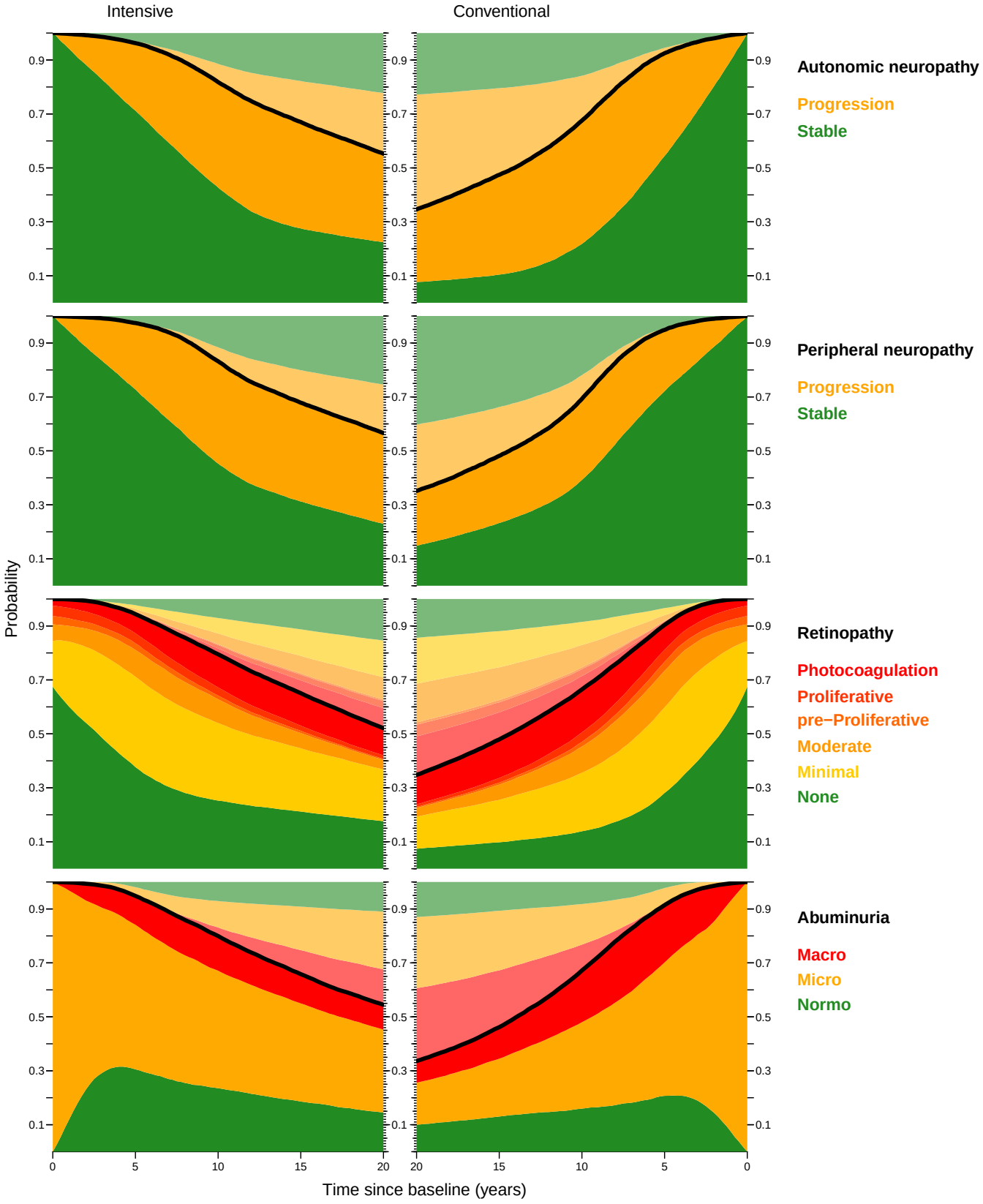

Figure 5.27: *Cumulative risks of microvascular complications, controlled for mortality. There is something fishy about the retinopathy — the overall mortality is too high.*

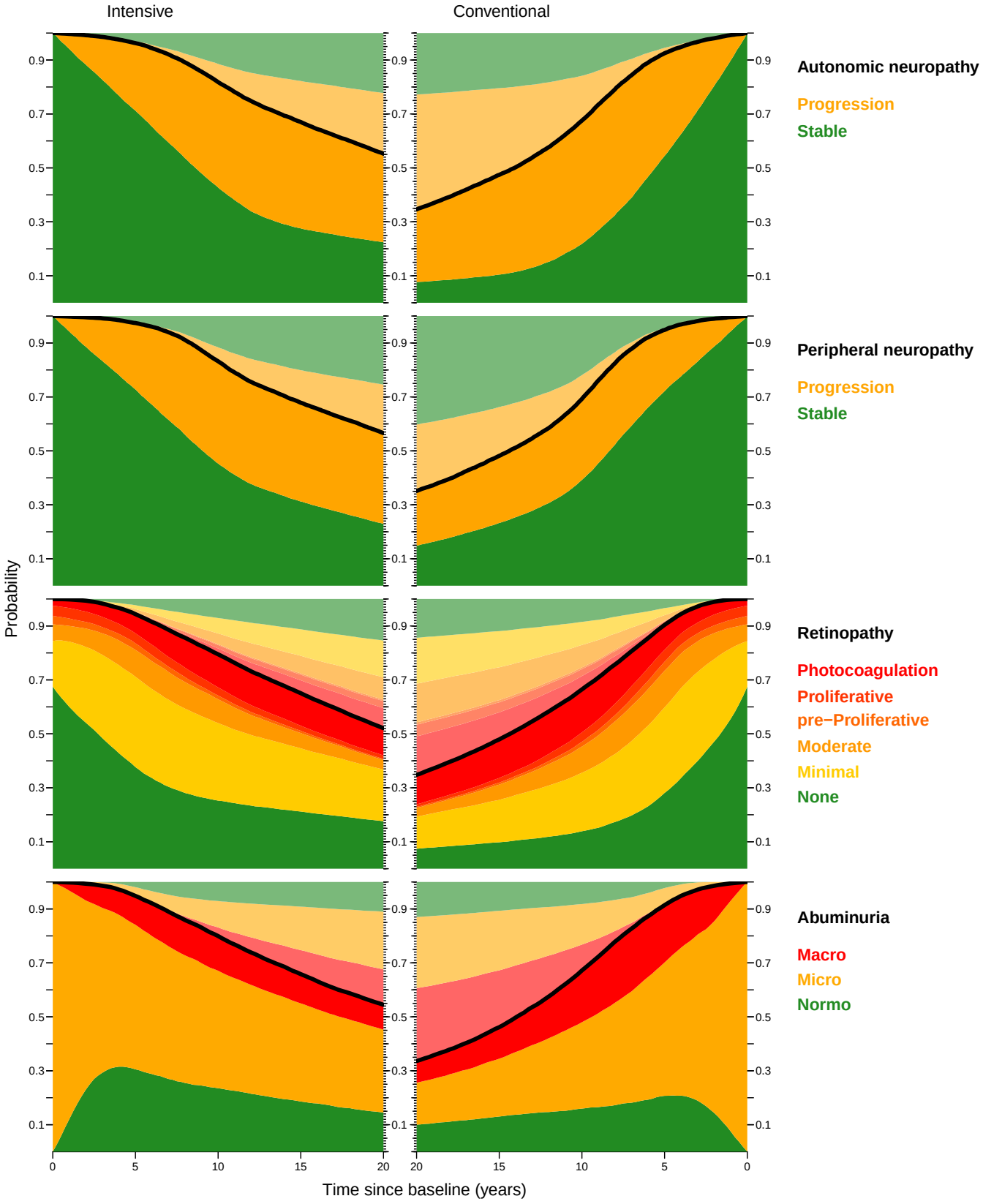

Figure 5.28: *Cumulative risks of microvascular complications, controlled for mortality. There is something fishy about the retinopathy — the overall mortality is too high.*

5.7 Plotting of summary HRs

```
> dimnames( mainCI )

$outcome
 [1] "All cause mortality"    "CVD mortality"          "non-CVD mortality"
 [4] "Death or 1st CVD"       "Death | CVD state"      "CVD event | CVD state"
 [7] "Retinopthy progression" "Autonomic neuropathy"   "Peripheral neuropathy"
[10] "Macroalbuminuria"

$model
[1] "raw"     "age/sex"

$what
[1] "Estimate" "2.5%"     "97.5%"    "P"
```

```
> yy <- 13-c(1,3,4,6,7,9:12)
>  plotEst( mainCI[-4,"raw",-4], xlog=TRUE, y=yy+0.1, txtpos=yy,
+           vref=1, grid=c(2:10,15,20)/10, xlim=c(0.2,2), lwd=3, cex=1.5,
+           xlab="Hazard ratio - Intensive versus Conventional")
```

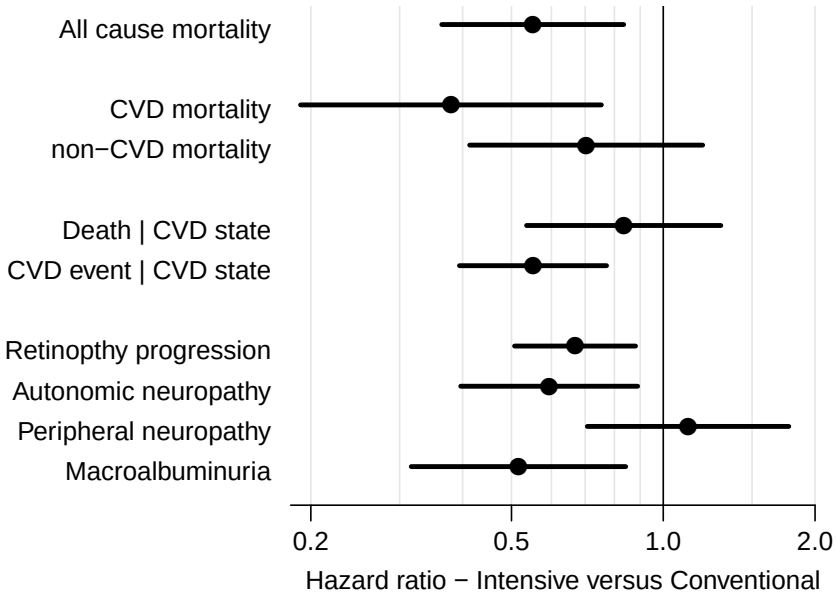

Figure 5.29: *HRs for different outcomes*

```
> fc <- function( x, d=2 ) formatC(x,digits=d,format="f")
> dimnames( mainCI )

$outcome
 [1] "All cause mortality"    "CVD mortality"          "non-CVD mortality"
 [4] "Death or 1st CVD"       "Death | CVD state"      "CVD event | CVD state"
 [7] "Retinopthy progression" "Autonomic neuropathy"   "Peripheral neuropathy"
[10] "Macroalbuminuria"

$model
[1] "raw"     "age/sex"

$what
[1] "Estimate" "2.5%"     "97.5%"    "P"
```

```
> yy <- 13-c(1,3,4,6,7,9:12)
>  plotEst( mainCI[-4,"raw",-4], xlog=TRUE, y=yy+0.1, txtpos=yy,
+           vref=1, grid=c(2:10,15,20)/10, xlim=c(0.2,5), lwd=3, cex=1.5,
+           xlab="Hazard ratio - Intensive versus Conventional")
> for( i in 1:length(yy) )
+ text( 2.0, yy[i],
+       paste( fc( mainCI[-4,,][i,"raw",1]), " (",
+              fc( mainCI[-4,,][i,"raw",2]), "-",
+              fc( mainCI[-4,,][i,"raw",3]), "), P=",
+              fc( mainCI[-4,,][i,"raw",4], 3),
+              sep = "" ),
+       adj = 0, cex=0.8 )
```

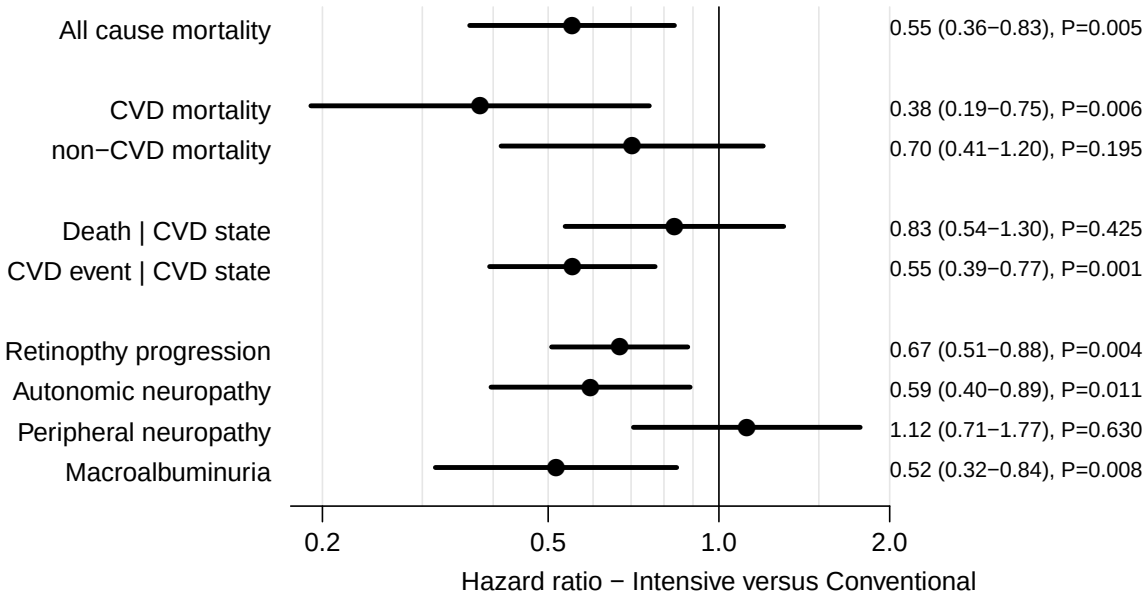

Figure 5.30: *HRs for different outcomes*

References

[1] P. Gaede, H. Lund-Andersen, H. H. Parving, and O. Pedersen. Effect of a multifactorial intervention on mortality in type 2 diabetes. *N. Engl. J. Med.*, 358(6):580–591, Feb 2008.