# Years of Life Lost (YLL) to disease Diabetes in DK as example

SDC
February 2017
http://bendixcarstensen.com/Epi
Version 1.2

Compiled Friday 14th April, 2017, 09:19
from: /home/bendix/stat/R/lib.src/Epi/pkg/vignettes/yll.tex

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

# Contents

# Chapter 1

# Theory and technicalities

This vignette for the `Epi` package describes the probabilistic/demographic background for and technical implementation of the `erl` and `yll` functions that computes the expected residual life time and years of life lost in an illness-death model.

## 1.1 Years of life lost (YLL)

...to diabetes or any other disease for that matter.

The general concept in calculation of "years lost to..." is the comparison of the expected lifetime between two groups of persons; one with and one without disease (in this example DM). The expected lifetime is the area under the survival curve, so basically the exercise requires that two survival curves that are deemed relevant be available.

The years of life lost is therefore just the area between the survival curves for those "Well", $S_W(t)$, and for those "Diseased", $S_D(t)$:

$$\text{YLL} = \int_0^\infty S_W(t) - S_D(t)\,\mathrm{d}t$$

The time $t$ could of course be age, but it could also be "time after age 50" and the survival curves compared would then be survival curves *conditional* on survival till age 50, and the YLL would be the years of life lost for a 50-year old person with diabetes.

If we are referring to the expected lifetime we will more precisely use the label expected residual lifetime, ERL.

## 1.2 Constructing the survival curves

YLL can be computed in two different ways, depending on the way the survival curve and hence the expected lifetime of a person *without* diabetes is computed:

- Assume that the "Well" persons are *immune* to disease — using only the non-DM mortality rates throughout for calculation of expected life time.

- Assume that the "Well" persons *can* acquire the disease and thereby see an increased mortality, thus involving all three rates shown in figure 1.1.

The former gives a higher YLL because the comparison is to persons assumed immune to DM (and yet with the same mortality as non-immune prior to diagnosis), the latter gives a more realistic picture of the comparison of group of persons with and without diabetes at a given age that can be interpreted in the real world.

The differences can be illustrated by figure 1.1; the immune approach corresponds to an assumption of $\lambda(t) = 0$ in the calculation of the survival curve for a person in the "Well" state.

Calculation of the survival of a diseased person already in the "DM" state is unaffected by assumptions about $\lambda$.

```
R version 3.3.3 (2017-03-06)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 14.04.5 LTS

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

other attached packages:
[1] Epi_2.10

loaded via a namespace (and not attached):
 [1] cmprsk_2.2-7      MASS_7.3-45       Matrix_1.2-6      plyr_1.8.4
 [5] parallel_3.3.3    survival_2.41-3   etm_0.6-2         Rcpp_0.12.5
 [9] splines_3.3.3     grid_3.3.3        numDeriv_2014.2-1 lattice_0.20-33
```

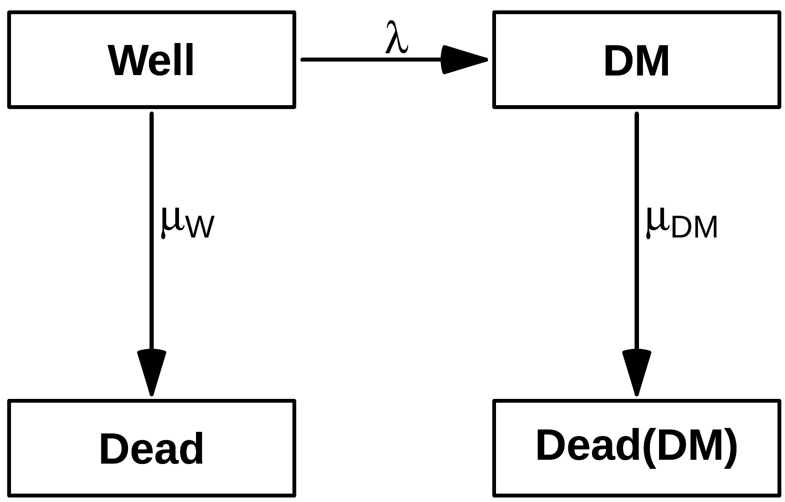


Figure 1.1: *Illness-death model describing diabetes incidence and -mortality.*

### 1.2.1 Total mortality — a shortcut?

A practical crude shortcut could be to compare the ERL in the diabetic population to the ERL for the *entire* population (that is use the total mortality ignoring diabetes status).

Note however that this approach also counts the mortality of persons that acquired the disease earlier, thus making the comparison population on average more ill than the population we aim at, namely those well at a given time, which only then become more gradually ill.

How large these effects are can however be empirically explored, as we shall do later.

### 1.2.2 Disease duration

In the exposition above there is no explicit provision for the effect of disease duration, but if we were able to devise mortality rates for any combination of age and duration, this could be taken into account.

There are however severe limitations in this as we in principle would want to have duration effects as long as the age-effects — in principle for all $(a, d)$ where $d \leq A$, where $A$ is the age at which we condition. So even if we were only to compute ERL from age, say, 40 we would still need duration effects up to 60 years (namely to age 100).

The incorporation of duration effects is in principle trivial from a computational point of view, but we would be forced to entertain models predicting duration effects way beyond what is actually observed disease duration in any practical case.

### 1.2.3 Computing integrals

The practical calculations of survival curves, ERL and YLL involves calculation of (cumulative) integrals of rates and functions of these as we shall see below. This is easy if we have a closed form expression of the function, so its value may be computed at any time point — this will be the case if we model rates by smooth parametric functions.

Computing the (cumulative) integral of a function is done as follows:

- Compute the value of the function (mortality rate for example) at the midpoints of a sequence of narrow equidistant intervals — for example one- or three month intervals of age, say.

- Take the cumulative sum of these values multiplied by the interval length — this will be a very close approximation to the cumulative integral evaluated at the end of each interval.

- If the intervals are really small (like 1/100 year), the distinction between the value at the middle and at the end of each interval becomes irrelevant.

Note that in the above it is assumed that the rates are given in units corresponding to the interval length — or more precisely, as the cumulative rates over the interval.

## 1.3 Survival functions in the illness-death model

The survival functions for persons in the "Well" state can be computed under two fundamentally different scenarios, depending on whether persons in the "Well" state are assumed to be immune to the disease ($\lambda(a) = 0$) or not.

### 1.3.1 Immune approach

In this case both survival functions for person in the two states are the usual simple transformation of the cumulative mortality rates:

$$S_W(a) = \exp\left(-\int_0^a \mu_W(u)\,\mathrm{d}u\right), \qquad S_D(a) = \exp\left(-\int_0^a \mu_D(u)\,\mathrm{d}u\right)$$

#### 1.3.1.1 Conditional survival functions

If we want the *conditional* survival functions given survival to age $A$, say, they are just:

$$S_W(a|A) = S_W(a)/S_W(A), \qquad S_D(a|A) = S_D(a)/S_D(A)$$

### 1.3.2 Non-immune approach

For a diseased person, the survival function in this states is the same as above, but the survival function for a person without disease (at age 0) is (see figure 1.1):

$$S(a) = \mathrm{P}\left\{\text{Well}\right\}(a) + \mathrm{P}\left\{\text{DM}\right\}(a)$$

In the appendix of the paper [2] is an indication of how to compute the probability of being in any of the four states shown in figure 1.1, which I shall repeat here:

In terms of the rates, the probability of being in the "Well" box is simply the probability of escaping both death (at a rate of $\mu_W(a)$) and diabetes (at a rate of $\lambda(a)$):

$$\mathrm{P}\left\{\text{Well}\right\}(a) = \exp\left(-\int_0^a \mu_W(u) + \lambda(u)\right)\,\mathrm{d}u$$

The probability of being alive with diabetes at age $a$, is computed given that diabetes occurred at age $s$ ($s < a$) and then integrated over $s$ from 0 to $a$:

$$\begin{aligned}
\mathrm{P}\left\{\text{DM}\right\}(a) &= \int_0^a \mathrm{P}\left\{\text{survive to } s\text{, DM diagnosed at } s\right\} \\
&\qquad \times \mathrm{P}\left\{\text{survive with DM from } s \text{ to } a\right\}\,\mathrm{d}s \\
&= \int_0^a \lambda(s)\exp\left(-\int_0^s \mu_W(u) + \lambda(u)\,\mathrm{d}u\right) \\
&\qquad \times \exp\left(-\int_s^a \mu_D(u)\,\mathrm{d}u\right)\,\mathrm{d}s
\end{aligned}$$

Sometimes we will use a version where the mortality among diabetes patients depend both on age $a$ and duration of diabetes, $d$, $\mu_D(a,d)$, in which case we get:

$$\begin{aligned}
\mathrm{P}\left\{\text{DM}\right\}(a) = \int_0^a &\lambda(s)\exp\left(-\int_0^s \mu_W(u) + \lambda(u)\,\mathrm{d}u\right) \\
&\times \exp\left(-\int_s^a \mu_D(u, u-s)\,\mathrm{d}u\right)\,\mathrm{d}s
\end{aligned}$$

because the integration variable $u$ is the age-scale and the second integral refers to mortality among persons diagnosed at age $s$, that is, with duration $u - s$ at age $u$.

The option of using duration-dependent mortality rates among diseased individuals is not implemented yet.

#### 1.3.2.1 Conditional survival functions

Unlike the immune approach, the conditional survival function in the more realistic case is not just a ratio of the unconditional to the value at the conditioning age, $A$, say. This would amount to conditioning on being merely *alive* at age $A$, but what we want is to condition on being in the "Well" state at age $A$.

The formulae for the conditional probabilities of being either in "Well" or "DM", given being in "Well" at age $A$ are basically replicates of the unconditional, albeit with changes in integration limits:

$$
\begin{aligned}
\mathrm{P}\left\{\text{Well}|\text{Well at } A\right\}(a) &= \exp\left(-\int_A^a \mu_W(u) + \lambda(u)\right)\,\mathrm{d}u \\
\mathrm{P}\left\{\text{DM}|\text{Well at } A\right\}(a) &= \int_A^a \lambda(s) \exp\left(-\int_A^s \mu_W(u) + \lambda(u)\,\mathrm{d}u\right) \\
&\qquad \times \exp\left(-\int_s^a \mu_D(u, u-s)\,\mathrm{d}u\right)\,\mathrm{d}s
\end{aligned}
$$

The calculation of these conditional survival functions is implemented but not allowing for duration-dependence. Thus it is only implemented assuming $\mu_D(a, d) = \mu_D(a)$.

# Chapter 2

# Analyses for DM in Denmark

The rates we use as basis for the following calculations are derived from the NDR, where we have omitted the blood-glucose criteria, because there is compelling evidence that these have quite a low specificity (particularly in the younger ages among women), and do not substantially contribute to the sensitivity.

As noted above the calculations of YLL requires access to (age-specific) rates of incidence of DM and mortality for persons with and without DM.

## 2.1 Modeling mortality and incidence data

We read in the dataset of DM and population mortality and incidence, `DMepi`:

```
> data( DMepi )
```

The dataset `DMepi` contains diabetes events, deaths and person-years for persons without diabetes and deaths and person-years for persons with diabetes:

```
> str( DMepi )
'data.frame':        4000 obs. of  8 variables:
 $ sex : Factor w/ 2 levels "M","F": 1 2 1 2 1 2 1 2 1 2 ...
 $ A   : num  0 0 1 1 2 2 3 3 4 4 ...
 $ P   : num  1996 1996 1996 1996 1996 ...
 $ X   : num  1 9 4 7 7 2 6 5 9 4 ...
 $ D.nD: num  28 19 23 19 7 8 8 8 6 7 ...
 $ Y.nD: num  35454 33095 36451 34790 35329 ...
 $ D.DM: num  0 0 0 0 0 0 0 0 0 0 ...
 $ Y.DM: num  0.476 3.877 4.92 7.248 12.474 ...
> head( DMepi )
  sex A    P X D.nD     Y.nD D.DM       Y.DM
1   M 0 1996 1   28 35453.65    0  0.4757016
2   F 0 1996 9   19 33094.86    0  3.8767967
3   M 1 1996 4   23 36450.73    0  4.9199179
4   F 1 1996 7   19 34789.99    0  7.2484600
5   M 2 1996 7    7 35328.92    0 12.4743326
6   F 2 1996 2    8 33673.43    0  8.0951403
```

For each combination of sex, age, period and date of birth in 1 year age groups, we have the person-years in the “Well” (`Y.nD`) and the “DM” (`Y.DM`) states, as well as the number of deaths from these (`D.nD`, `D.DM`) and the number of incident diabetes cases from the “Well” state (`X`).

In order to compute the years of life lost to diabetes and how this has changed over time, we fit models for the mortality and incidence of both groups (and of course, separately for men and women). The models we use will be age-period-cohort models [1] providing estimated mortality rates for ages 0–99 and dates 1.1.1996–1.1.2016.

First we transform the age and period variables to reflect the mean age and period in each of the Lexis triangles. We also compute the total number of deaths and amount of risk time, as we are going to model the total mortality as well. Finally we restrict the dataset to ages over 30 only:

```
> DMepi <- transform( subset( DMepi, A>30 ),
+                     D.T = D.nD + D.DM,
+                     Y.T = Y.nD + Y.DM )
> head(DMepi)
   sex  A    P  X D.nD     Y.nD D.DM     Y.DM D.T      Y.T
63   M 31 1996 21   51 43909.32    0 291.4107  51 44200.73
64   F 31 1996 33   16 41376.91    2 287.4969  18 41664.41
65   M 32 1996 26   67 43159.94    0 299.6571  67 43459.59
66   F 32 1996 20   23 40706.49    1 275.2615  24 40981.75
67   M 33 1996 35   54 41251.06    4 321.0397  58 41572.10
68   F 33 1996 32   23 39102.29    1 277.4463  24 39379.74
```

With the correct age and period coding in the Lexis triangles, we fit models for the mortalities and incidences. Note that we for comparative purposes also fit a model for the *total* mortality, ignoring the

```
> # Knots used in all models
> ( a.kn <- seq(40,95,,6) )
[1] 40 51 62 73 84 95
> ( p.kn <- seq(1997,2015,,4) )
[1] 1997 2003 2009 2015
> ( c.kn <- seq(1910,1976,,6) )
[1] 1910.0 1923.2 1936.4 1949.6 1962.8 1976.0
> # Check the number of events between knots
> ae <- xtabs( cbind(D.nD,D.DM,X) ~ cut(A,c(30,a.kn,Inf)) + sex, data=DMepi )
> ftable( addmargins(ae,1), col.vars=3:2 )
                                D.nD          D.DM            X       
                        sex        M      F      M      F      M      F
cut(A, c(30, a.kn, Inf))                                                
(30,40]                         8899   4525    564    269   9912   8622
(40,51]                        24686  15296   2886   1399  31668  20769
(51,62]                        57747  38968  10276   4916  53803  34495
(62,73]                       102877  78771  24070  13008  51000  38731
(73,84]                       154804 153842  31006  25414  25444  26804
(84,95]                        97698 175484  13972  21231   4726   7852
(95,Inf]                        5800  20563    545   1522     74    238
Sum                           452511 487449  83319  67759 176627 137511
> pe <- xtabs( cbind(D.nD,D.DM,X) ~ cut(P,c(1990,p.kn,Inf)) + sex, data=DMepi )
> ftable( addmargins(pe,1), col.vars=3:2 )
                                  D.nD          D.DM            X       
                          sex        M      F      M      F      M      F
cut(P, c(1990, p.kn, Inf))                                                
(1990,1997]                      51901  54162   6012   5378  12477  10030
(1997,2003]                     145418 157768  21028  17976  43749  34255
```

```
(2003,2009]                    133175 144717  25172  20595  56556  43891
(2009,2015]                    122017 130802  31107  23810  63845  49335
(2015,Inf]                          0      0      0      0      0      0
Sum                            452511 487449  83319  67759 176627 137511
> ce <- xtabs( cbind(D.nD,D.DM,X) ~ cut(P-A,c(-Inf,c.kn,Inf)) + sex, data=DMepi )
> ftable( addmargins(ce,1), col.vars=3:2 )
                                    D.nD          D.DM            X       
                               sex     M      F      M      F      M      F
cut(P - A, c(-Inf, c.kn, Inf))                                            
(-Inf,1.91e+03]                    19912  49797   1784   3731    536   1143
(1.91e+03,1.92e+03]               130691 190012  18160  23709   9765  13519
(1.92e+03,1.94e+03]               154227 146284  32435  24876  34897  32481
(1.94e+03,1.95e+03]                93397  67909  22921  11437  65012  44292
(1.95e+03,1.96e+03]                40948  26234   6724   3326  46155  29891
(1.96e+03,1.98e+03]                12534   6810   1245    654  19293  15311
(1.98e+03, Inf]                      802    403     50     26    969    874
Sum                               452511 487449  83319  67759 176627 137511
> # Fit an APC-model for all transitions, seperately for men and women
> mW.m <- glm( D.nD ~ -1 + Ns(A   ,knots=a.kn,int=TRUE) +
+                          Ns(  P,knots=p.kn,ref=2005) +
+                          Ns(P-A,knots=c.kn,ref=1950),
+             offset = log(Y.nD),
+             family = poisson,
+               data = subset( DMepi, sex=="M" ) )
> mD.m <- update( mW.m,  D.DM ~ . , offset=log(Y.DM) )
> mT.m <- update( mW.m,  D.T  ~ . , offset=log(Y.T ) )
> lW.m <- update( mW.m,  X ~ . )
> # Model for women
> mW.f <- update( mW.m, data = subset( DMepi, sex=="F" ) )
> mD.f <- update( mD.m, data = subset( DMepi, sex=="F" ) )
> mT.f <- update( mT.m, data = subset( DMepi, sex=="F" ) )
> lW.f <- update( lW.m, data = subset( DMepi, sex=="F" ) )
```

## 2.2 Residual life time and years lost to DM

We now collect the estimated years of life lost classified by method (immune assumption or not), sex, age and calendar time:

```
> a.ref <- 30:90
> p.ref <- 1996:2016
> aYLL <- NArray( list( type = c("Imm","Tot","Sus"),
+                        sex = levels( DMepi$sex ),
+                        age = a.ref,
+                       date = p.ref ) )
> str( aYLL )
 logi [1:3, 1:2, 1:61, 1:21] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 4
  ..$ type: chr [1:3] "Imm" "Tot" "Sus"
  ..$ sex : chr [1:2] "M" "F"
  ..$ age : chr [1:61] "30" "31" "32" "33" ...
  ..$ date: chr [1:21] "1996" "1997" "1998" "1999" ...
> system.time(
+ for( ip in p.ref )
+    {
```

```
+     nd <- data.frame( A = seq(30,90,0.2)+0.1,
+                       P = ip,
+                    Y.nD = 1,
+                    Y.DM = 1,
+                    Y.T  = 1 )
+     muW.m <- ci.pred( mW.m, nd )[,1]
+     muD.m <- ci.pred( mD.m, nd )[,1]
+     muT.m <- ci.pred( mT.m, nd )[,1]
+     lam.m <- ci.pred( lW.m, nd )[,1]
+     muW.f <- ci.pred( mW.f, nd )[,1]
+     muD.f <- ci.pred( mD.f, nd )[,1]
+     muT.f <- ci.pred( mT.f, nd )[,1]
+     lam.f <- ci.pred( lW.f, nd )[,1]
+     aYLL["Imm","M",,paste(ip)] <- yll( int=0.2, muW.m, muD.m, lam=NULL,
+                                        A=a.ref, age.in=30, note=FALSE )[-1]
+     aYLL["Imm","F",,paste(ip)] <- yll( int=0.2, muW.f, muD.f, lam=NULL,
+                                        A=a.ref, age.in=30, note=FALSE )[-1]
+     aYLL["Tot","M",,paste(ip)] <- yll( int=0.2, muT.m, muD.m, lam=NULL,
+                                        A=a.ref, age.in=30, note=FALSE )[-1]
+     aYLL["Tot","F",,paste(ip)] <- yll( int=0.2, muT.f, muD.f, lam=NULL,
+                                        A=a.ref, age.in=30, note=FALSE )[-1]
+     aYLL["Sus","M",,paste(ip)] <- yll( int=0.2, muW.m, muD.m, lam=lam.m,
+                                        A=a.ref, age.in=30, note=FALSE )[-1]
+     aYLL["Sus","F",,paste(ip)] <- yll( int=0.2, muW.f, muD.f, lam=lam.f,
+                                        A=a.ref, age.in=30, note=FALSE )[-1]
+     } )
   user  system elapsed
 16.951   0.000  16.955
> round( ftable( aYLL[,,seq(1,61,10),], col.vars=c(3,2) ), 1 )
          age   30         40         50         60         70         80         90
          sex    M    F     M    F     M    F     M    F     M    F     M    F     M    F
type date
Imm  1996       11.7 10.8  9.8  9.7  7.8  8.2  5.6  6.3  3.4  3.9  1.5  1.6  0.0  0.0
     1997       11.5 10.6  9.7  9.4  7.7  7.9  5.6  6.1  3.4  3.9  1.4  1.6  0.0  0.0
     1998       11.3 10.3  9.6  9.2  7.6  7.7  5.5  5.9  3.4  3.8  1.4  1.5  0.0  0.0
     1999       11.1 10.0  9.4  9.0  7.5  7.5  5.4  5.7  3.3  3.7  1.4  1.5  0.0  0.0
     2000       10.9  9.8  9.3  8.7  7.4  7.3  5.4  5.6  3.3  3.6  1.4  1.5  0.0  0.0
     2001       10.7  9.5  9.1  8.5  7.3  7.1  5.3  5.4  3.3  3.4  1.4  1.4  0.0  0.0
     2002       10.5  9.2  9.0  8.3  7.1  6.9  5.2  5.2  3.2  3.3  1.3  1.4  0.0  0.0
     2003       10.3  9.0  8.8  8.1  7.0  6.7  5.1  5.1  3.1  3.2  1.3  1.3  0.0  0.0
     2004       10.0  8.8  8.6  7.8  6.8  6.5  5.0  4.9  3.1  3.1  1.3  1.3  0.0  0.0
     2005        9.7  8.5  8.4  7.6  6.6  6.3  4.8  4.8  3.0  3.0  1.2  1.3  0.0  0.0
     2006        9.4  8.3  8.1  7.5  6.5  6.2  4.7  4.6  2.9  2.9  1.2  1.2  0.0  0.0
     2007        9.1  8.1  7.9  7.3  6.3  6.0  4.6  4.5  2.8  2.8  1.1  1.2  0.0  0.0
     2008        8.9  7.9  7.7  7.1  6.1  5.9  4.4  4.3  2.7  2.7  1.1  1.1  0.0  0.0
     2009        8.6  7.7  7.5  6.9  6.0  5.7  4.3  4.2  2.7  2.6  1.1  1.1  0.0  0.0
     2010        8.4  7.5  7.3  6.8  5.9  5.6  4.2  4.1  2.6  2.5  1.1  1.1  0.0  0.0
     2011        8.3  7.3  7.2  6.7  5.8  5.5  4.2  4.0  2.6  2.5  1.0  1.0  0.0  0.0
     2012        8.1  7.2  7.1  6.5  5.7  5.4  4.1  4.0  2.6  2.4  1.0  1.0  0.0  0.0
     2013        8.0  7.1  7.0  6.4  5.6  5.3  4.1  3.9  2.5  2.4  1.0  1.0  0.0  0.0
     2014        7.8  6.9  6.9  6.3  5.6  5.3  4.1  3.8  2.5  2.3  1.0  0.9  0.0  0.0
     2015        7.7  6.8  6.8  6.2  5.5  5.2  4.0  3.8  2.5  2.2  1.0  0.9  0.0  0.0
     2016        7.6  6.7  6.7  6.1  5.5  5.1  4.0  3.7  2.5  2.2  1.0  0.9  0.0  0.0
Tot  1996       11.1 10.4  9.3  9.2  7.3  7.7  5.2  5.9  3.1  3.7  1.3  1.5  0.0  0.0
     1997       10.9 10.1  9.1  9.0  7.2  7.5  5.1  5.7  3.1  3.6  1.3  1.5  0.0  0.0
     1998       10.7  9.8  9.0  8.7  7.0  7.3  5.0  5.5  3.1  3.5  1.3  1.4  0.0  0.0
     1999       10.5  9.6  8.8  8.5  6.9  7.1  5.0  5.4  3.0  3.4  1.3  1.4  0.0  0.0
     2000       10.3  9.3  8.6  8.3  6.8  6.9  4.9  5.2  3.0  3.3  1.3  1.4  0.0  0.0
```

```
     2001     10.0  9.0  8.5  8.0  6.6  6.6  4.8  5.0  2.9  3.2  1.2  1.3  0.0  0.0
     2002      9.8  8.8  8.3  7.8  6.5  6.4  4.7  4.8  2.9  3.1  1.2  1.3  0.0  0.0
     2003      9.5  8.5  8.1  7.6  6.3  6.2  4.5  4.7  2.8  2.9  1.2  1.2  0.0  0.0
     2004      9.3  8.2  7.8  7.3  6.1  6.0  4.4  4.5  2.7  2.8  1.1  1.2  0.0  0.0
     2005      9.0  8.0  7.6  7.1  5.9  5.9  4.3  4.3  2.6  2.7  1.1  1.1  0.0  0.0
     2006      8.7  7.8  7.4  6.9  5.8  5.7  4.1  4.2  2.5  2.6  1.0  1.1  0.0  0.0
     2007      8.4  7.5  7.1  6.7  5.6  5.5  4.0  4.1  2.4  2.5  1.0  1.1  0.0  0.0
     2008      8.1  7.3  6.9  6.6  5.4  5.4  3.8  3.9  2.3  2.4  1.0  1.0  0.0  0.0
     2009      7.8  7.1  6.7  6.4  5.2  5.2  3.7  3.8  2.3  2.3  0.9  1.0  0.0  0.0
     2010      7.6  7.0  6.5  6.3  5.1  5.1  3.6  3.7  2.2  2.2  0.9  0.9  0.0  0.0
     2011      7.4  6.8  6.4  6.1  5.0  5.0  3.5  3.6  2.2  2.2  0.9  0.9  0.0  0.0
     2012      7.3  6.6  6.3  6.0  4.9  4.9  3.5  3.5  2.1  2.1  0.9  0.9  0.0  0.0
     2013      7.1  6.5  6.1  5.9  4.8  4.8  3.4  3.4  2.1  2.0  0.9  0.8  0.0  0.0
     2014      7.0  6.3  6.0  5.8  4.8  4.7  3.4  3.4  2.0  2.0  0.8  0.8  0.0  0.0
     2015      6.8  6.2  5.9  5.6  4.7  4.6  3.3  3.3  2.0  1.9  0.8  0.8  0.0  0.0
     2016      6.7  6.1  5.8  5.5  4.6  4.6  3.3  3.2  2.0  1.9  0.8  0.7  0.0  0.0
Sus  1996     10.7 10.1  8.9  9.0  7.1  7.6  5.2  5.9  3.2  3.8  1.4  1.6  0.0  0.0
     1997     10.5  9.8  8.8  8.8  7.0  7.4  5.1  5.7  3.2  3.7  1.4  1.5  0.0  0.0
     1998     10.3  9.6  8.7  8.5  6.9  7.2  5.1  5.5  3.2  3.6  1.4  1.5  0.0  0.0
     1999     10.2  9.3  8.5  8.3  6.8  7.0  5.0  5.4  3.1  3.5  1.4  1.5  0.0  0.0
     2000     10.0  9.0  8.4  8.1  6.7  6.8  4.9  5.2  3.1  3.4  1.4  1.4  0.0  0.0
     2001      9.7  8.8  8.2  7.8  6.5  6.6  4.8  5.0  3.1  3.3  1.3  1.4  0.0  0.0
     2002      9.5  8.5  8.1  7.6  6.4  6.4  4.7  4.9  3.0  3.2  1.3  1.4  0.0  0.0
     2003      9.2  8.3  7.8  7.4  6.2  6.2  4.6  4.7  2.9  3.1  1.3  1.3  0.0  0.0
     2004      8.9  8.0  7.6  7.2  6.0  5.9  4.5  4.5  2.9  3.0  1.2  1.3  0.0  0.0
     2005      8.6  7.7  7.3  6.9  5.8  5.8  4.3  4.4  2.8  2.8  1.2  1.2  0.0  0.0
     2006      8.3  7.5  7.1  6.7  5.6  5.6  4.2  4.2  2.7  2.7  1.1  1.2  0.0  0.0
     2007      8.0  7.2  6.8  6.5  5.4  5.4  4.0  4.1  2.6  2.6  1.1  1.1  0.0  0.0
     2008      7.7  7.0  6.6  6.3  5.2  5.2  3.9  3.9  2.5  2.5  1.1  1.1  0.0  0.0
     2009      7.4  6.8  6.4  6.2  5.1  5.1  3.8  3.8  2.4  2.4  1.0  1.1  0.0  0.0
     2010      7.3  6.6  6.2  6.0  5.0  5.0  3.7  3.7  2.4  2.4  1.0  1.0  0.0  0.0
     2011      7.1  6.5  6.2  5.9  4.9  4.9  3.7  3.7  2.4  2.3  1.0  1.0  0.0  0.0
     2012      7.0  6.4  6.1  5.8  4.9  4.9  3.6  3.6  2.3  2.3  1.0  1.0  0.0  0.0
     2013      7.0  6.3  6.0  5.8  4.9  4.8  3.6  3.6  2.3  2.2  1.0  0.9  0.0  0.0
     2014      6.9  6.3  6.0  5.7  4.9  4.8  3.6  3.5  2.3  2.2  1.0  0.9  0.0  0.0
     2015      6.8  6.2  6.0  5.7  4.9  4.8  3.6  3.5  2.3  2.1  1.0  0.9  0.0  0.0
     2016      6.8  6.1  6.0  5.6  4.9  4.7  3.6  3.5  2.3  2.1  1.0  0.8  0.0  0.0
```

We now have the relevant points for the graph showing YLL to diabetes for men and women by age, and calendar year, both under the immunity and susceptibility models for the calculation of YLL.

```
> plyll <- function(wh){
+ par( mfrow=c(1,2), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n", las=1 )
+
+ matplot( a.ref, aYLL[wh,"M",,],
+          type="l", lty=1, col="blue", lwd=1:2,
+          ylim=c(0,12), xlab="Age",
+          ylab="Years lost to DM", yaxs="i" )
+ abline(v=50,h=1:10,col=gray(0.7))
+ text( 90, 11, "Men", col="blue", adj=1 )
+ text( 40, aYLL[wh,"M","40","1996"], "1996", adj=c(0,0), col="blue" )
+ text( 43, aYLL[wh,"M","44","2016"], "2016", adj=c(1,1), col="blue" )
+
+ matplot( a.ref, aYLL[wh,"F",,],
+          type="l", lty=1, col="red", lwd=1:2,
+          ylim=c(0,12), xlab="Age",
+          ylab="Years lost to DM", yaxs="i" )
```

```
+ abline(v=50,h=1:10,col=gray(0.7))
+ text( 90, 11, "Women", col="red", adj=1 )
+ text( 40, aYLL[wh,"F","40","1996"], "1996", adj=c(0,0), col="red" )
+ text( 43, aYLL[wh,"F","44","2016"], "2016", adj=c(1,1), col="red" )
+ }
> plyll("Imm")
```

```
> plyll("Tot")
```

```
> plyll("Sus")
```

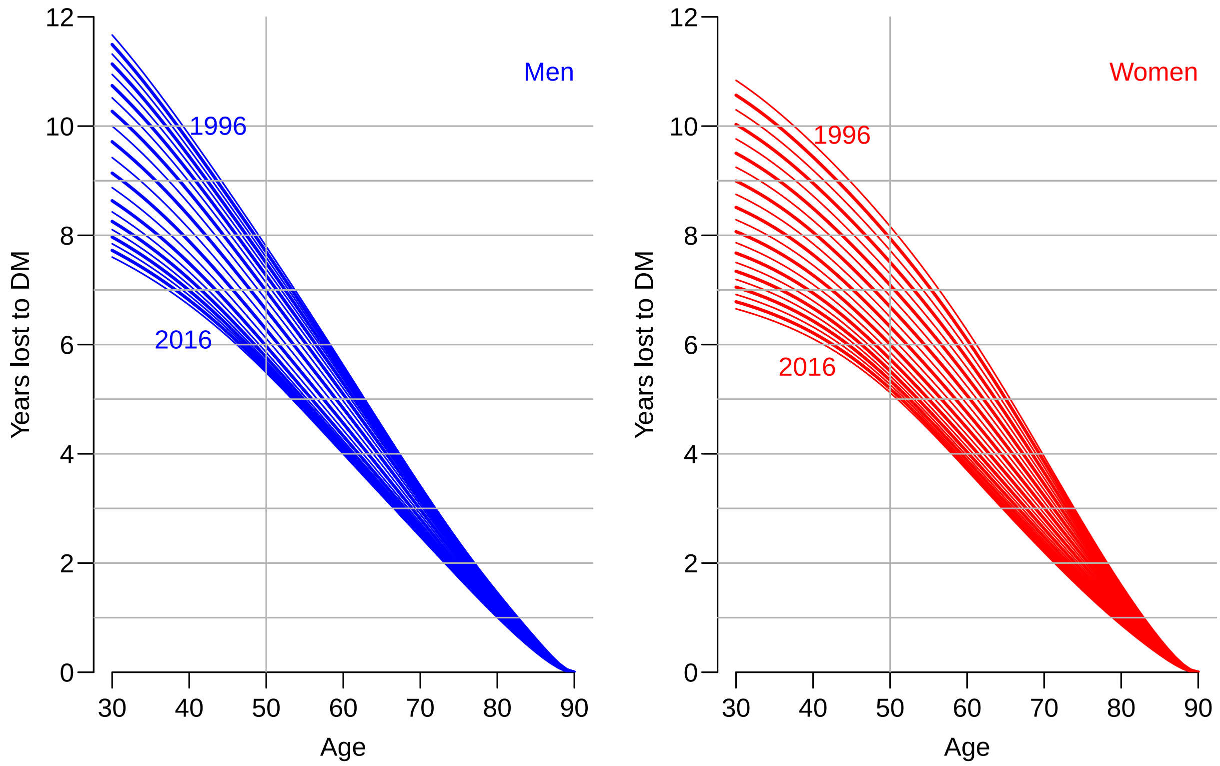


Figure 2.1: *Years of life lost to DM: the difference in expected residual life time at different ages between persons with and without diabetes, assuming the persons without diabetes at a given age remain free from diabetes (immunity assumption — not reasonable). The lines refer to date of evaluation; the top lines refer to 1.1.1996 the bottom ones to 1.1.2016. Blue curves are men, red women.*

From figure 2.2 we see that for men aged 50 the years lost to diabetes has decreased from a bit over 8 to a bit less than 6 years, and for women from 8.5 to 5 years; so a greater improvement for women.

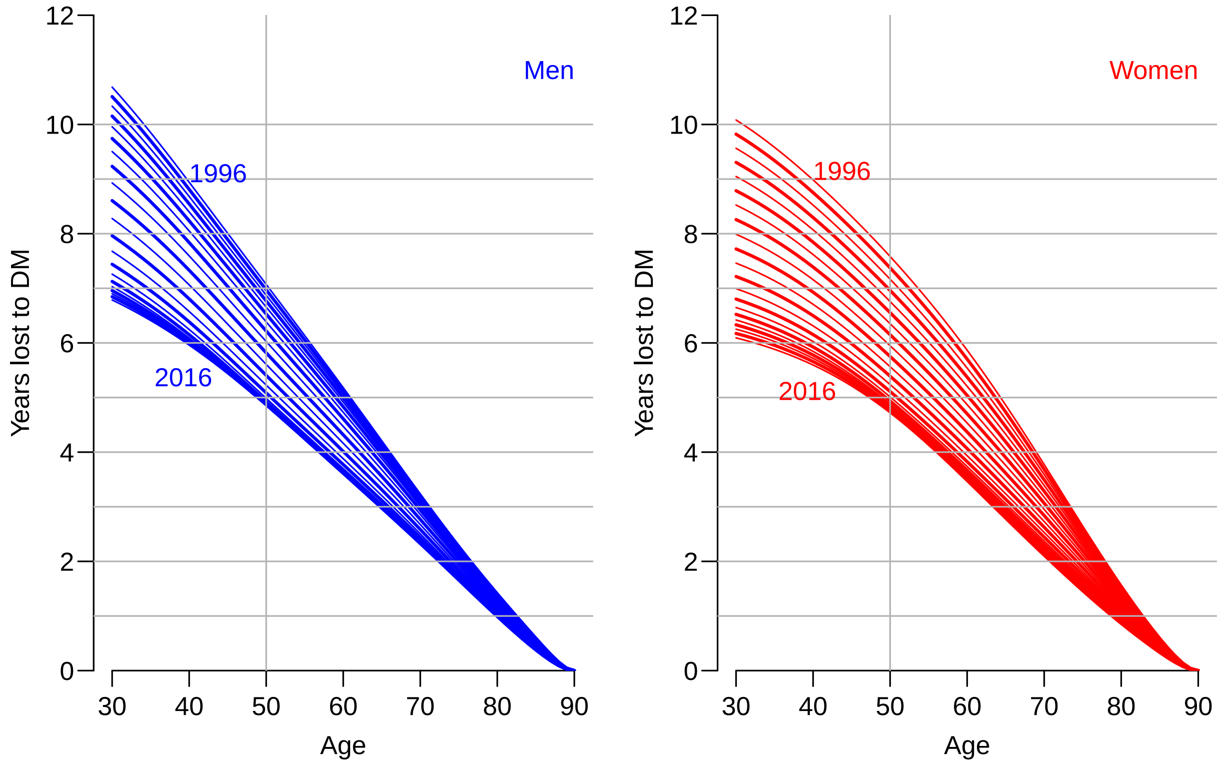


Figure 2.2: *Years of life lost to DM: the difference in expected residual life time at different ages between persons with and without diabetes, allowing the persons without diabetes at a given to contract diabetes and thus be subject to higher mortality. The lines refer to date of evaluation; the top lines refer to 1.1.1996 the bottom ones to 1.1.2016. Blue curves are men, red women.*

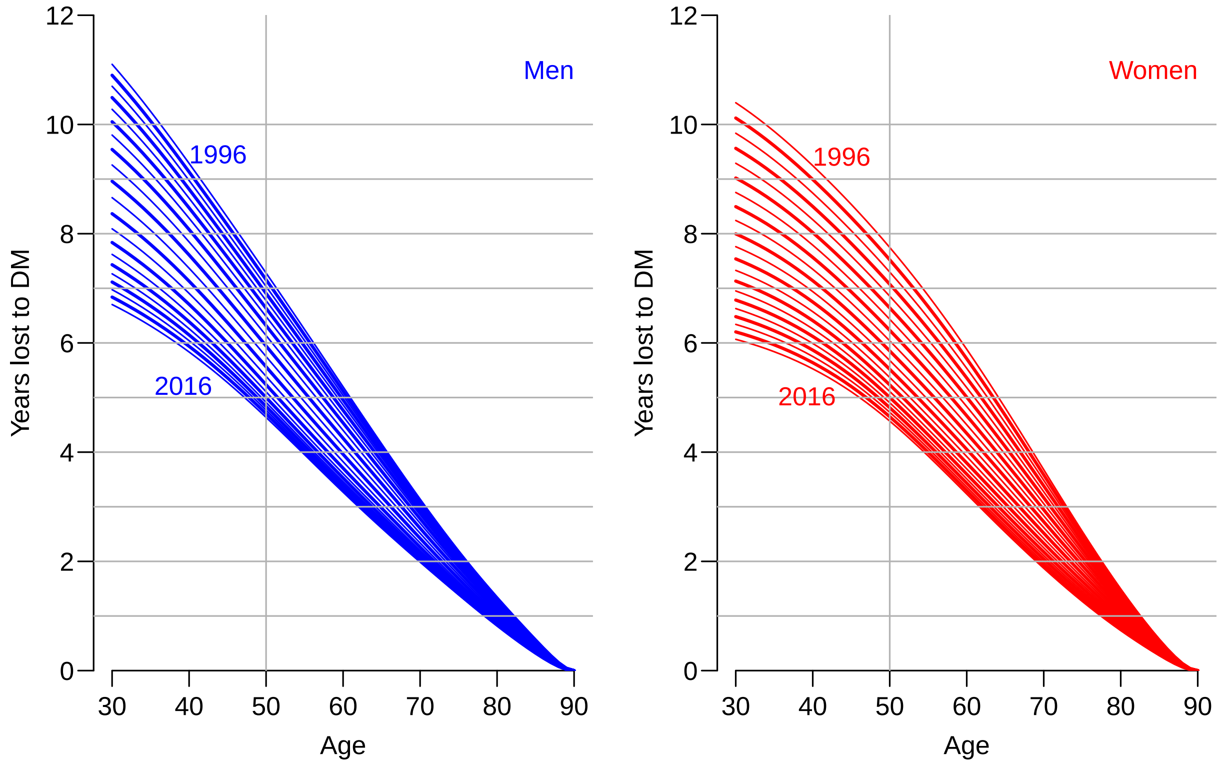


Figure 2.3: *Years of life lost to DM: the difference in expected residual life time at different ages between persons with and without diabetes. Allowance for susceptibility is approximated by using the total population mortality instead of non-DM mortality. The lines refer to date of evaluation; the top lines refer to 1.1.1996 the bottom ones to 1.1.2016. Blue curves are men, red women.*

# Chapter 3

# Practical implementation

We have devised functions that wraps these formulae up for practical use.

## 3.1 Function definitions

When using the functions it is assumed that the functions $\mu_W$, $\mu_D$ and $\lambda$ are given as vectors corresponding to equidistantly (usually tightly) spaced ages from 0 to $K$ where K is the age where everyone can safely be assumed dead.

`surv1` is a simple function that computes the survival function from a vector of mortality rates, and optionally the conditional survival given being alive at prespecified ages:

```
> surv1
function( int, mu, age.in=0, A=NULL )
{
# Computes the survival function from age A till the end, assuming
# that mu is a vector of mortalities in intervals of length int.
# int and mu should be in compatible units that is T and T^-1 for
# some unit T (months, years, ...)

# age-class boundaries
age <- 0:length(mu)*int + age.in

# cumulative rates and survival at the boundaries
Mu <- c( 0, cumsum( mu )*int )
Sv <- exp( -Mu )
surv <- data.frame( age=age, surv=Sv )

# if a vector of conditioning ages A is given
if( cond <- !is.null(A) )
  {
  j <- 0
  # actual conditioning ages
  cage <- NULL
  for( ia in A )
     {
     j <- j+1
     # Where is the age we condition on
     cA <- which( diff(age>ia)==1 )
     surv <- cbind( surv, pmin( 1, surv$surv/(surv$surv[cA]) ) )
     cage[j] <- surv$age[cA]
     }
```

```
  }
names( surv )[-1] <- paste( "A", c( age.in, if( cond ) cage else NULL ), sep="" )
rownames( surv ) <- NULL
return( surv )
}
```

`erl1` basically just expands the result of `surv1` with a column of expected residual life times:

```
> erl1
function( int, mu, age.in = 0 )
{
# Computes expected residual life time at all ages
age <- 0:length(mu)*int + age.in

# Small utility: cumulative cumulative sum from the end of a vector
musmuc <- function( x ) rev( cumsum( rev(x) ) )

# The survival function with a 0 at end, and the integral from the upper end
surv <- surv1( int = int, mu = mu, age.in = age.in )[,2]
cbind( age = age,
      surv = surv,
       erl = c( musmuc( ( surv[-1]-diff(surv)/2 ) ) /
                 surv[-length(surv)], 0 ) * int )
}
```

We also define a function, `surv2`, that computes the survival function for a non-diseased person that may become diseased with rate `lam` and after that die at a rate of `muD` (corresponding to the formulae above). This is the sane way of handling years of life lost to a particular illness:

```
> surv2
function( int, muW, muD, lam, age.in=0, A=NULL )
{
# check the vectors
if( length(muW) != length(muD) |
    length(muD) != length(lam) )
  stop( "Vectors with rates must have same length:\n",
          "length(muW)=", length(muW),
        ", length(muD)=", length(muD),
        ", length(lam)=", length(lam) )

# First the workhorse that computes the survival function for a
# person in Well assuming that the mortality rate from this state is
# muW, disease incidence is in lam, and mortality in the diseased
# state is muD, and that all refer to constant rates intervals of
# length int starting from age.in, conditional on survival to A
wsurv2 <-
function( int, muW, muD, lam, age.in=0, A=0 )
{
# age-class boundaries - note one longer that rate vectors refers to
# boundaries of intervals not midpoints
age <- 0:length(muW)*int + age.in

# cumulative rates at the boundaries, given survival to A
MuW <-  cumsum( c( 0, muW ) * ( age > A ) ) * int
MuD <-  cumsum( c( 0, muD ) * ( age > A ) ) * int
Lam <-  cumsum( c( 0, lam ) * ( age > A ) ) * int
```

```
# probability of being well
pW <- exp( -( Lam + MuW ) )

# probability of diagnosis at s --- first term in the integral for
# P(DM at a). Note that we explicitly add a 0 at the start so we get a
# probability of 0 of transition at the first age point
Dis <- c(0,lam) * ( age > A ) * exp( -(Lam+MuW) ) * int

# for each age (age[ia]) we compute the integral over the range
# [0,age] of the product of the probability of diagnosis and the
# probability of surviving from diagnosis till age ia
pDM <- Dis * 0
for( ia in 1:length(age) )
   pDM[ia] <- sum( Dis[1:ia] * exp( -(MuD[ia]-MuD[1:ia]) ) )
                  # 1st term as function of s (1:ia)
                              # 2nd term integral over range s:age
                  # upper integration limit is age (ia) and the lower
                  # limit is the intermediate age (at DM) (1:ia)
# Finally, we add the probabilities of being in Well resp. DM to get
# the overall survival:
surv <- data.frame( age = age, surv = pDM + pW )
return( surv )
}

# survival from start
surv <- wsurv2( int, muW, muD, lam, age.in=age.in, A=0 )

# add columns for conditioning ages
if( !is.null(A) )
   {
   for( j in 1:length(A) )
      {
      surv <- cbind( surv,
                     wsurv2( int, muW, muD, lam, age.in=age.in, A=A[j] )[,2] )
      }
   }
Al <- A
for( i in 1:length(A) ) Al[i] <- max( surv$age[surv$age <= A[i]] )
colnames( surv )[-1] <- paste( "A", c( age.in, Al ), sep="" )

# done!
return( surv )
}
```

Finally we devised a function using these to compute the expected residual lifetime at select ages:

```
> erl
function( int,
          muW,
          muD,
          lam = NULL,
       age.in = 0,
            A = NULL,
       immune = is.null(lam),
          yll = TRUE,
         note = TRUE )
{
```

```
# Computes expected residual life time for Well and Dis states
# respectively in an illness-death model, optionally ignoring
# the well->ill transition

# Utility to integrate a survival function from the last point where
# it is 1, assuming points are 1 apart
trsum <-
function( x )
{
x[c(diff(x)==0,TRUE)] <- NA
sum( ( x[-length(x)] + x[-1] ) / 2, na.rm=TRUE )
}

# Check sensibility
if( !immune & is.null(lam) ) stop( "'lam' is required when immune=FALSE\n" )

# Survival functions
             sD <- surv1( int=int,      muD,      age.in = age.in, A = A )
if( immune ) sW <- surv1( int=int, muW,           age.in = age.in, A = A )
else         sW <- surv2( int=int, muW, muD, lam, age.in = age.in, A = A )

# Area under the survival functions
erl <- cbind( apply( sW[,-1], 2, trsum ),
              apply( sD[,-1], 2, trsum ) ) * int
colnames( erl ) <- c("Well","Dis")
rownames( erl ) <- colnames( sW )[-1]

# Should we compute years of life lost?
if( yll ) erl <- cbind( erl, YLL = erl[,"Well"] - erl[,"Dis"] )

# Cautionary note
if( immune )
  {
  attr( erl, "NOTE" ) <- "Calculations assume that Well persons cannot get Ill (quite silly
  if( note ) cat("NOTE:",  attr( erl, "NOTE" ), "\n" )
  }
return( erl )
}
```

. . . and a wrapper for this if we only want the years of life lost returned:

```
> yll
function( int,
          muW,
          muD,
          lam = NULL,
       age.in = 0,
            A = NULL,
       immune = is.null(lam),
         note = TRUE ) erl( int = int,
                            muW = muW,
                            muD = muD,
                            lam = lam,
                         age.in = age.in,
                              A = A,
                         immune = immune,
                            yll = TRUE,
                           note = note )[,"YLL"]
```

# Bibliography


[1] B Carstensen. Age-Period-Cohort models for the Lexis diagram. *Statistics in Medicine*, 26(15):3018–3045, July 2007.

[2] B. Carstensen, J.K. Kristensen, P. Ottosen, and K. Borch-Johnsen. The Danish National Diabetes Register: Trends in incidence, prevalence and mortality. *Diabetologia*, 51:2187–2196, 2008.