A note on step sizes in calculation and simulation in multistate models

SDC (at CDC)
http://bendixcarstensen.com/
June 2015
Version 1

Compiled Monday 15th June, 2015, 19:10
from: /home/bendix/sdc/DMreg/DemoDMCa/steps.tex

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

Contents

1 Introduction

When simulating transitions in a multistate model, it is standard practice to simulate based on only the probabilities of a single transition from a given state and ignore the probability that further transitions occur in the interval used. This introduces of course some error; the larger the larger the simulation interval. However, it is not really clear how large this error is in practical circumstances.

To illustrate the magnitude of the possible errors we here consider the transition probabilities from a study of diabetes and cancer occurrence and use these actual probabilities to sketch the size of the problem.

2 Example multistate model

We use as example a demographic study on diabetes and cancer, based on estimating the transition intensities in the multistate model shown in figure 1. We have estimated all 9

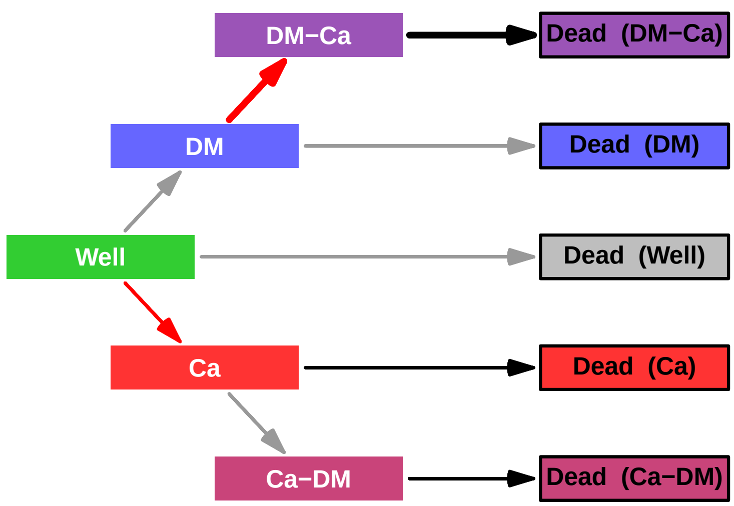

Figure 1: *Multistate model of diabetes and cancer occurrence in Denmark 1995–2012.*

transition rates as a smooth function of age, calendar time and date of birth, separately for men and women. predictions for each 1 January and ages 0–100 in steps of 1 month are in the PR array:

```
> library( Epi )
> load( file="./data/rates.Rda")
> str( PR )
```

```
 num [1:10, 1:10, 1:1224, 1:18, 1:2] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 5
  ..$ from: chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
```

```
  ..$ to  : chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ age : chr [1:1224] "0.0416666666666667" "0.125" "0.208333333333333" "0.291666666666667" ...
  ..$ per : chr [1:18] "1995" "1996" "1997" "1998" ...
  ..$ sex : chr [1:2] "M" "F"
```

We shall use the rates between ages 70 and 75 to illustrate our point and to facilitate things we extract the slice we shall work with:

```
> round( as.numeric( dimnames(PR)[["age"]][70*12+0:61] ), 3 )
 [1] 69.958 70.042 70.125 70.208 70.292 70.375 70.458 70.542 70.625 70.708 70.792 70.875 70.958
[14] 71.042 71.125 71.208 71.292 71.375 71.458 71.542 71.625 71.708 71.792 71.875 71.958 72.042
[27] 72.125 72.208 72.292 72.375 72.458 72.542 72.625 72.708 72.792 72.875 72.958 73.042 73.125
[40] 73.208 73.292 73.375 73.458 73.542 73.625 73.708 73.792 73.875 73.958 74.042 74.125 74.208
[53] 74.292 74.375 74.458 74.542 74.625 74.708 74.792 74.875 74.958 75.042
```

```
> PR <- PR[,,70*12+1:60,"2005","M"]
> str( PR )
 num [1:10, 1:10, 1:60] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 3
  ..$ from: chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ to  : chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ age : chr [1:60] "70.0416666666667" "70.125" "70.2083333333333" "70.2916666666667" ...
```

```
> round( PR[,,1] * 10000 )
        to
from     Well DM DM-Ca Ca Ca-DM D-W D-DM D-Ca D-DC D-CD
  Well      0 13     0 20     0  10    0    0    0    0
  DM        0  0    24  0     0   0   29    0    0    0
  DM-Ca     0  0     0  0     0   0    0    0  252    0
  Ca        0  0     0  0    17   0    0  116    0    0
  Ca-DM     0  0     0  0     0   0    0    0    0   97
  D-W       0  0     0  0     0   0    0    0    0    0
  D-DM      0  0     0  0     0   0    0    0    0    0
  D-Ca      0  0     0  0     0   0    0    0    0    0
  D-DC      0  0     0  0     0   0    0    0    0    0
  D-CD      0  0     0  0     0   0    0    0    0    0
```

The third dimension of the array is labeled by the midpoint of the interval that the 1-month integrated intensities refer to.

Thus we have the matrices of transition *intensities* or rater the 1-month cumulative intensities, but we need to transform these to 1-month transition probabilities, so that we also have the probabilities of remaining in each state in the diagonal. To this end we devise a small function, `ci2pr`, that converts a matrix with off-diagonal elements of cumulative intensities

```
> ci2pr <-
+ function( M )
+ {
+ sm  <- apply( M, 1, sum )
+ res <- sweep( M, 1, (1-exp(-sm))/sm, "*" )
+ # Rows corresponding to absorbing states have sum 0 so the above
+ # returns NA, which must then be converted to 0 before the diagonal is
+ # filled with the survival probabilities
+ res[is.na(res)] <- 0
+ diag( res ) <- exp( -sm )
+ res
+ }
> round( ci2pr( PR[,,1] )*10000 )
```

```
       to
from    Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD
  Well  9957   13     0   20     0    10     0     0     0     0
  DM       0 9947    24    0     0     0    29     0     0     0
  DM-Ca    0    0  9751    0     0     0     0     0   249     0
  Ca       0    0     0 9868    17     0     0   116     0     0
  Ca-DM    0    0     0    0  9903     0     0     0     0    97
  D-W      0    0     0    0     0 10000     0     0     0     0
  D-DM     0    0     0    0     0     0 10000     0     0     0
  D-Ca     0    0     0    0     0     0     0 10000     0     0
  D-DC     0    0     0    0     0     0     0     0 10000     0
  D-CD     0    0     0    0     0     0     0     0     0 10000
```

But we have a whole sequence of matrices with cumulative intensities in `PR`, so we convert them all to one-month probability transition matrices:

```
> PRm <- apply( PR, 3, ci2pr )
> str( PRm )
 num [1:100, 1:60] 0.996 0 0 0 0 ...
 - attr(*, "dimnames")=List of 2
  ..$    : NULL
  ..$ age: chr [1:60] "70.0416666666667" "70.125" "70.2083333333333" "70.2916666666667" ...
> # Note: apply kills the first two dimensions, sor reinsitute them:
> dim( PRm ) <- dim( PR )
> dimnames( PRm ) <- dimnames( PR )
> round( addmargins( PRm[,,31,drop=FALSE]*10000, 2 ) )
, , age = 72.5416666666667

       to
from    Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well  9948   14     0   25     0    14     0     0     0     0 10000
  DM       0 9938    27    0     0     0    35     0     0     0 10000
  DM-Ca    0    0  9753    0     0     0     0     0   247     0 10000
  Ca       0    0     0 9856    16     0     0   128     0     0 10000
  Ca-DM    0    0     0    0  9894     0     0     0     0   106 10000
  D-W      0    0     0    0     0 10000     0     0     0     0 10000
  D-DM     0    0     0    0     0     0 10000     0     0     0 10000
  D-Ca     0    0     0    0     0     0     0 10000     0     0 10000
  D-DC     0    0     0    0     0     0     0     0 10000     0 10000
  D-CD     0    0     0    0     0     0     0     0     0 10000 10000
```

Thus `PRm` now contains the 1-month transition probabilities under the assumption that two transitions do not occur in the same month. This is of course not correct but quite an accurate approximation. We shall empirically investigate how computed 5-year transition probabilities depend on the choice of interval length (or more specifically on the size of the integrated intensities) for this example.

3 The 5 year transition probabilities

Specifically, we are considering the transition probabilities between states from age exactly 70 to age exactly 75 years.

We shall look at how the transition probabilities would look if we based them on the cumulative intensities for 3-month, 6-month, one-year intervals and finally 5-year intervals.

For the intervals with an even number of months, we use the average of the two middle intensities, corresponding to the scaled incidence rates at the midpoint of the 6-month, respectively 1-year intervals.

3.1 Using 1 month probabilities

Now in order to get the 5-year probability transition matrix using the 1 month cumulative intensities we must multiply the matrices with transition probabilities:

```
> p5.1m <- PRm[,,1]
> for( i in 2:60) p5.1m <- PRm[,,i] %*% p5.1m
> round( addmargins( p5.1m*10000, 2 ) )
```

```
       to
from     Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well   7306  600    33  813    42   722    67   388    20    10 10000
  DM        0 6867   653    0     0     0  1783     0   697     0 10000
  DM-Ca     0    0  2146    0     0     0     0     0  7854     0 10000
  Ca        0    0     0 4182   463     0     0  5174     0   181 10000
  Ca-DM     0    0     0    0  5242     0     0     0     0  4758 10000
  D-W       0    0     0    0     0 10000     0     0     0     0 10000
  D-DM      0    0     0    0     0     0 10000     0     0     0 10000
  D-Ca      0    0     0    0     0     0     0 10000     0     0 10000
  D-DC      0    0     0    0     0     0     0     0 10000     0 10000
  D-CD      0    0     0    0     0     0     0     0     0 10000 10000
```

3.2 Using 3 month probabilities

Using instead 3 month probabilities we first compute the cumulative incidence rates over 3 months intervals by taking the cumulative rate in the middle and multiply with 3:

```
> as.numeric( dimnames(PR)[["age"]][0:19*3+2] )
```

```
 [1] 70.125 70.375 70.625 70.875 71.125 71.375 71.625 71.875 72.125 72.375 72.625 72.875 73.125
[14] 73.375 73.625 73.875 74.125 74.375 74.625 74.875
```

```
> Ptmp <- PR[,,0:19*3+2]*3
> str( Ptmp )
```

```
 num [1:10, 1:10, 1:20] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 3
  ..$ from: chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ to  : chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ age : chr [1:20] "70.125" "70.375" "70.625" "70.875" ...
```

```
> PRq <- apply( Ptmp, 3, ci2pr )
> dim( PRq ) <- dim( Ptmp )
> dimnames( PRq ) <- dimnames( Ptmp )
> str( PRq )
```

```
 num [1:10, 1:10, 1:20] 0.987 0 0 0 0 ...
 - attr(*, "dimnames")=List of 3
  ..$ from: chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ to  : chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ age : chr [1:20] "70.125" "70.375" "70.625" "70.875" ...
```

```
> round( PRq[,,5]*10000 )
```

```
      to
from    Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD
  Well  9861   41     0   65     0    33     0     0     0     0
  DM       0 9829    76    0     0     0    94     0     0     0
  DM-Ca    0    0  9266    0     0     0     0     0   734     0
  Ca       0    0     0 9594    49     0     0   357     0     0
  Ca-DM    0    0     0    0  9702     0     0     0     0   298
  D-W      0    0     0    0     0 10000     0     0     0     0
  D-DM     0    0     0    0     0     0 10000     0     0     0
  D-Ca     0    0     0    0     0     0     0 10000     0     0
  D-DC     0    0     0    0     0     0     0     0 10000     0
  D-CD     0    0     0    0     0     0     0     0     0 10000
```

As before, we compute the 5-year transition probabilities between states based on the half-year cumulative intensities:

```
> p5.3m <- PRq[,,1]
> for( i in 2:20) p5.3m <- PRq[,,i] %*% p5.3m
> round( addmargins( p5.3m*10000, 2 ) )
      to
from    Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well  7306  604    33  825    41   722    65   378    18     9 10000
  DM       0 6867   670    0     0     0  1783     0   680     0 10000
  DM-Ca    0    0  2146    0     0     0     0     0  7854     0 10000
  Ca       0    0     0 4182   468     0     0  5174     0   176 10000
  Ca-DM    0    0     0    0  5242     0     0     0     0  4758 10000
  D-W      0    0     0    0     0 10000     0     0     0     0 10000
  D-DM     0    0     0    0     0     0 10000     0     0     0 10000
  D-Ca     0    0     0    0     0     0     0 10000     0     0 10000
  D-DC     0    0     0    0     0     0     0     0 10000     0 10000
  D-CD     0    0     0    0     0     0     0     0     0 10000 10000
```

3.3 Using 6 month probabilities

Using instead 6 month probabilities we first compute the cumulative incidence rates over 6 months intervals by averaging:

```
> round( as.numeric( dimnames(PR)[["age"]][0:9*6+3] ), 3 )
 [1] 70.208 70.708 71.208 71.708 72.208 72.708 73.208 73.708 74.208 74.708
> round( as.numeric( dimnames(PR)[["age"]][0:9*6+4] ), 3 )
 [1] 70.292 70.792 71.292 71.792 72.292 72.792 73.292 73.792 74.292 74.792
> Ptmp <- (PR[,,0:9*6+3]+PR[,,0:9*6+3])/2*6
> dimnames(Ptmp)[["age"]] <- round( ( as.numeric( dimnames(PR)[["age"]][0:9*6+3] ) +
+                                     as.numeric( dimnames(PR)[["age"]][0:9*6+4] ) )/2, 2 )
> str( Ptmp )
 num [1:10, 1:10, 1:10] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 3
  ..$ from: chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ to  : chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ age : chr [1:10] "70.25" "70.75" "71.25" "71.75" ...
> PRh <- apply( Ptmp, 3, ci2pr )
> dim( PRh ) <- dim( Ptmp )
> dimnames( PRh ) <- dimnames( Ptmp )
> str( PRh )
```

```
 num [1:10, 1:10, 1:10] 0.974 0 0 0 0 ...
 - attr(*, "dimnames")=List of 3
  ..$ from: chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ to  : chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ age : chr [1:10] "70.25" "70.75" "71.25" "71.75" ...
```

```
> round( PRh[,,5]*10000 )
       to
from     Well    DM DM-Ca    Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD
  Well   9699    83     0   142     0    77     0     0     0     0
  DM        0  9641   156     0     0     0   203     0     0     0
  DM-Ca     0     0  8605     0     0     0     0     0  1395     0
  Ca        0     0     0  9173    96     0     0   730     0     0
  Ca-DM     0     0     0     0  9387     0     0     0     0   613
  D-W       0     0     0     0     0 10000     0     0     0     0
  D-DM      0     0     0     0     0     0 10000     0     0     0
  D-Ca      0     0     0     0     0     0     0 10000     0     0
  D-DC      0     0     0     0     0     0     0     0 10000     0
  D-CD      0     0     0     0     0     0     0     0     0 10000
```

As before, we compute the 5-year transition probabilities between states based on the half-year cumulative intensities:

```
> p5.6m <- PRh[,,1]
> for( i in 2:10) p5.6m <- PRh[,,i] %*% p5.6m
> round( addmargins( p5.6m*10000, 2 ) )
       to
from     Well    DM DM-Ca    Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well   7313   610    33   841    40   718    62   360    16     8 10000
  DM        0  6874   695     0     0     0  1777     0   653     0 10000
  DM-Ca     0     0  2149     0     0     0     0     0  7851     0 10000
  Ca        0     0     0  4187   477     0     0  5167     0   169 10000
  Ca-DM     0     0     0     0  5248     0     0     0     0  4752 10000
  D-W       0     0     0     0     0 10000     0     0     0     0 10000
  D-DM      0     0     0     0     0     0 10000     0     0     0 10000
  D-Ca      0     0     0     0     0     0     0 10000     0     0 10000
  D-DC      0     0     0     0     0     0     0     0 10000     0 10000
  D-CD      0     0     0     0     0     0     0     0     0 10000 10000
```

3.4 Using 1 year probabilities

Similar to the 6-month case we do the calculations using 1-year cumulative rates:

```
> round( as.numeric( dimnames(PR)[["age"]][0:4*12+6] ), 3 )
[1] 70.458 71.458 72.458 73.458 74.458
> round( as.numeric( dimnames(PR)[["age"]][0:4*12+7] ), 3 )
[1] 70.542 71.542 72.542 73.542 74.542
> Ptmp <- (PR[,,0:4*12+6]+PR[,,0:4*12+7])/2*12
> dimnames(Ptmp)[["age"]] <- round( as.numeric( dimnames(Ptmp)[["age"]] ), 1 )
> str(Ptmp)
 num [1:10, 1:10, 1:5] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "dimnames")=List of 3
  ..$ from: chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ to  : chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ age : chr [1:5] "70.5" "71.5" "72.5" "73.5" ...
```

```
> PR1 <- apply( Ptmp, 3, ci2pr )
> dim( PR1 ) <- dim( Ptmp )
> dimnames( PR1 ) <- dimnames( Ptmp )
> str( PR1 )

 num [1:10, 1:10, 1:5] 0.948 0 0 0 0 ...
 - attr(*, "dimnames")=List of 3
  ..$ from: chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ to  : chr [1:10] "Well" "DM" "DM-Ca" "Ca" ...
  ..$ age : chr [1:5] "70.5" "71.5" "72.5" "73.5" ...

> addmargins( round( PR1[,,3]*10000 ), 2 )

       to
from     Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well   9394  164     0  285     0   157     0     0     0     0 10000
  DM        0 9283   309    0     0     0   408     0     0     0 10000
  DM-Ca     0    0  7409    0     0     0     0     0  2591     0 10000
  Ca        0    0     0 8400   183     0     0  1417     0     0 10000
  Ca-DM     0    0     0    0  8797     0     0     0     0  1203 10000
  D-W       0    0     0    0     0 10000     0     0     0     0 10000
  D-DM      0    0     0    0     0     0 10000     0     0     0 10000
  D-Ca      0    0     0    0     0     0     0 10000     0     0 10000
  D-DC      0    0     0    0     0     0     0     0 10000     0 10000
  D-CD      0    0     0    0     0     0     0     0     0 10000 10000
```

As before, we compute the 5-year transition probabilities between states based on the 1-year cumulative intensities:

```
> p5.1y <- PR1[,,1]
> for( i in 2:5) p5.1y <- PR1[,,i] %*% p5.1y
> round( addmargins( p5.1y*10000, 2 ) )
       to
from     Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well   7306  621    32  880    39   721    55   328    11     5 10000
  DM        0 6867   751    0     0     0  1782     0   600     0 10000
  DM-Ca     0    0  2147    0     0     0     0     0  7853     0 10000
  Ca        0    0     0 4182   492     0     0  5173     0   153 10000
  Ca-DM     0    0     0    0  5243     0     0     0     0  4757 10000
  D-W       0    0     0    0     0 10000     0     0     0     0 10000
  D-DM      0    0     0    0     0     0 10000     0     0     0 10000
  D-Ca      0    0     0    0     0     0     0 10000     0     0 10000
  D-DC      0    0     0    0     0     0     0     0 10000     0 10000
  D-CD      0    0     0    0     0     0     0     0     0 10000 10000
```

3.5 Using 5 year probabilities

Finally for a very coarse comparison we also use the 5-year cumulative rates

```
> round( as.numeric( dimnames(PR)[["age"]][31:32] ), 3 )
[1] 72.542 72.625

> PR5 <- (PR[,,31]+PR[,,32])/2*12*5
> p5.5y <- ci2pr( PR5 )
> round( p5.5y*10000 )
```

```
      to
from     Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD
  Well   7299  726     0 1271     0   705     0     0     0     0
  DM        0 6880  1342    0     0     0  1777     0     0     0
  DM-Ca     0    0  2232    0     0     0     0     0  7768     0
  Ca        0    0     0 4172   664     0     0  5165     0     0
  Ca-DM     0    0     0    0  5256     0     0     0     0  4744
  D-W       0    0     0    0     0 10000     0     0     0     0
  D-DM      0    0     0    0     0     0 10000     0     0     0
  D-Ca      0    0     0    0     0     0     0 10000     0     0
  D-DC      0    0     0    0     0     0     0     0 10000     0
  D-CD      0    0     0    0     0     0     0     0     0 10000
```

3.6 Comparing the results

The comparison is simply to print the resulting 70 to 75 year transition probabilities. The lower half of the matrix is just the identity so we do not print it for brevity:

```
> print.table( addmargins( round( p5.1m[1:5,]*10000 ), 2 ), z="." )
      to
from     Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well   7306  600    33  813    42   722    67   388    20    10 10001
  DM        . 6867   653    .     .     .  1783     .   697     . 10000
  DM-Ca     .    .  2146    .     .     .     .     .  7854     . 10000
  Ca        .    .     . 4182   463     .     .  5174     .   181 10000
  Ca-DM     .    .     .    .  5242     .     .     .     .  4758 10000
> print.table( addmargins( round( p5.3m[1:5,]*10000 ), 2 ), z="." )
      to
from     Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well   7306  604    33  825    41   722    65   378    18     9 10001
  DM        . 6867   670    .     .     .  1783     .   680     . 10000
  DM-Ca     .    .  2146    .     .     .     .     .  7854     . 10000
  Ca        .    .     . 4182   468     .     .  5174     .   176 10000
  Ca-DM     .    .     .    .  5242     .     .     .     .  4758 10000
> print.table( addmargins( round( p5.6m[1:5,]*10000 ), 2 ), z="." )
      to
from     Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well   7313  610    33  841    40   718    62   360    16     8 10001
  DM        . 6874   695    .     .     .  1777     .   653     .  9999
  DM-Ca     .    .  2149    .     .     .     .     .  7851     . 10000
  Ca        .    .     . 4187   477     .     .  5167     .   169 10000
  Ca-DM     .    .     .    .  5248     .     .     .     .  4752 10000
> print.table( addmargins( round( p5.1y[1:5,]*10000 ), 2 ), z="." )
      to
from     Well   DM DM-Ca   Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well   7306  621    32  880    39   721    55   328    11     5  9998
  DM        . 6867   751    .     .     .  1782     .   600     . 10000
  DM-Ca     .    .  2147    .     .     .     .     .  7853     . 10000
  Ca        .    .     . 4182   492     .     .  5173     .   153 10000
  Ca-DM     .    .     .    .  5243     .     .     .     .  4757 10000
> print.table( addmargins( round( p5.5y[1:5,]*10000 ), 2 ), z="." )
```

```
      to
from     Well    DM DM-Ca    Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD   Sum
  Well   7299   726     .  1271     .   705     .     .     .     . 10001
  DM        .  6880  1342     .     .     .  1777     .     .     .  9999
  DM-Ca     .     .  2232     .     .     .     .     .  7768     . 10000
  Ca        .     .     .  4172   664     .     .  5165     .     . 10001
  Ca-DM     .     .     .     .  5256     .     .     .     .  4744 10000
```

Of equal interest of course also to see the successive differences between the approaches:

```
> print.table( addmargins( round( (p5.3m-p5.1m)[1:5,]*10000 ), 2 ), z="." )
      to
from    Well  DM DM-Ca  Ca Ca-DM D-W D-DM D-Ca D-DC D-CD Sum
  Well     .   4     .  12     .   .   -2  -11   -2   -1   .
  DM       .   .    17   .     .   .    .    .  -17    .   .
  DM-Ca    .   .     .   .     .   .    .    .    .    .   .
  Ca       .   .     .   .     5   .    .    .    .   -5   .
  Ca-DM    .   .     .   .     .   .    .    .    .    .   .
> print.table( addmargins( round( (p5.6m-p5.1m)[1:5,]*10000 ), 2 ), z="." )
      to
from    Well  DM DM-Ca  Ca Ca-DM D-W D-DM D-Ca D-DC D-CD Sum
  Well     7  10     .  29    -1  -4   -6  -28   -4   -2   1
  DM       .   7    42   .     .   .   -6    .  -44    .  -1
  DM-Ca    .   .     3   .     .   .    .    .   -3    .   .
  Ca       .   .     .   5    14   .    .   -6    .  -13   .
  Ca-DM    .   .     .   .     6   .    .    .    .   -6   .
> print.table( addmargins( round( (p5.1y-p5.1m)[1:5,]*10000 ), 2 ), z="." )
      to
from    Well  DM DM-Ca  Ca Ca-DM D-W D-DM D-Ca D-DC D-CD Sum
  Well     .  21     .  68    -3  -1  -12  -60   -9   -4   .
  DM       .   1    98   .     .   .   -1    .  -97    .   1
  DM-Ca    .   .     1   .     .   .    .    .   -1    .   .
  Ca       .   .     .   .    29   .    .   -1    .  -29  -1
  Ca-DM    .   .     .   .     1   .    .    .    .   -1   .
> print.table( addmargins( round( (p5.5y-p5.1m)[1:5,]*10000 ), 2 ), z="." )
      to
from    Well   DM DM-Ca   Ca Ca-DM  D-W D-DM D-Ca D-DC D-CD  Sum
  Well    -7  125   -33  458   -42  -17  -67 -388  -20  -10   -1
  DM       .   13   689    .     .    .   -5    . -697    .    .
  DM-Ca    .    .    87    .     .    .    .    .  -87    .    .
  Ca       .    .     .  -10   201    .    .   -9    . -181    1
  Ca-DM    .    .     .    .    14    .    .    .    .  -14    .
```

4 Conclusion

It is seen that there is fairly little difference between the 1 month and 3 month approach, indicating that both will provide a good approximation of the ultimately correct calculation based on infinitesimally small intervals.

Already using 6-month intervals starts being inaccurate. But note that this is not a feature of the interval length, but rater of the cumulative rates over the intervals, so we should be wary when these are of a magnitude as seen for the 6-month intervals or more, here we show the values of the half-year cumulative incidence rates in units of 1/10,000:

```
> print.table( round( PRh[,,5]*10000 ), z="." )
       to
from     Well    DM DM-Ca    Ca Ca-DM   D-W  D-DM  D-Ca  D-DC  D-CD
  Well   9699    83     .   142     .    77     .     .     .     .
  DM        .  9641   156     .     .     .   203     .     .     .
  DM-Ca     .     .  8605     .     .     .     .     .  1395     .
  Ca        .     .     .  9173    96     .     .   730     .     .
  Ca-DM     .     .     .     .  9387     .     .     .     .   613
  D-W       .     .     .     .     . 10000     .     .     .     .
  D-DM      .     .     .     .     .     . 10000     .     .     .
  D-Ca      .     .     .     .     .     .     . 10000     .     .
  D-DC      .     .     .     .     .     .     .     . 10000     .
  D-CD      .     .     .     .     .     .     .     .     . 10000
```

Thus it appears that a bold rule of thumb is that if you have cumulative rates exceeding 2% (200 per 10,000), you are presumably on the verge of getting inaccuracies in your calculations.