Sample size and precision in the Greenland complications study

SDC / MaEJ
http://bendixcarstensen.com/SDC/Greenland
October 2014
1

Compiled Wednesday 29th July, 2015, 16:22
from: /home/bendix/sdc/coll/maej/grCompl/pwrprc

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

Contents

1 The study

The Greenland complications study is an extra follow-up of the B99 and the IHIT cohorts, where persons with diabetes or impaired glucose regulation (IGR) are followed up for microvascular complications, notably nephropathy and retinopathy.

The number of persons in the IGR is about 1200 (not known till the 2014 survey is completed), with an estimated follow-up time of 7,940 person-years *after* onset of DM/IGR.

2 Outcomes

We want to estimate:

- rates of microvascular complications
- rate-ratios of microvascular complications between carriers of different genotypes of interest.

2.1 Assumptions

Based on previous studies and sound clinical judgment the rate of newly onset nephropathy among persons with DM or IGR is taken to be 20 per 1000 PY and the corresponding rate of retinopathy to be 30 per 1000 PY.

3 Precision and power

3.1 Rates

A bold calculation yields that a complication with a rate of 20/1000 during 7,940 PY will produce $20 \times 7.94 = 159$ events, and one with rates 30/1000 will produce 238 events:

```
> c(20,30) * 7.94
[1] 158.8 238.2
```

The usual formula for the relative precision of a rate estimate based on these numbers is:

$$\text{erf} = \exp(1.96/\sqrt{D})$$

which in the cases here will produce rate estimates with a relative precision in the vicinity of ± 15%:

```
> exp( 1.96/sqrt(c(159,238)) )
[1] 1.168170 1.135471
```

3.2 Rate-ratios

We are looking at genotypes with population prevalences between 5 and 10% separately and 44% in total, but we do not have any specific estimates of the likely risk elevation associated each of them.

Hence we will compute the precision of the RR estimates using a baseline rates of 20 and 30 per 1000 PY, for exposure prevalences 5 – 50 %, and exposure effects 1.05, . . . , 1.50. This will slightly overestimate the precision if the true overall rates is 20 resp. 30, because a fraction of the population having an elevated risk will of course result in an overall rate that is higher than the rate in the unexposed group. For the sake of completeness we also make the computations for person-years between 8,000 and 13,000.

Thus, we set up an array to collect the median RR (for control), the precision (in terms of the error factor, the relative uncertainty of the RR estimates) and the power for combinations of these factors:

```
> library( Epi )
> nsim <- 10000
> prpw <- NArray( list( RR = seq(1.05,1.5,0.05),
+                       pr = seq(5,50,5),
+                       PY = 8:13,
+                      inc = c(20,30),
+                     what = c("RR","Erf","Pwr") ) )
> str( prpw )
 logi [1:10, 1:10, 1:6, 1:2, 1:3] NA NA NA NA NA NA ...
 - attr(*, "dimnames")=List of 5
  ..$ RR  : chr [1:10] "1.05" "1.1" "1.15" "1.2" ...
  ..$ pr  : chr [1:10] "5" "10" "15" "20" ...
  ..$ PY  : chr [1:6] "8" "9" "10" "11" ...
  ..$ inc : chr [1:2] "20" "30"
  ..$ what: chr [1:3] "RR" "Erf" "Pwr"
```

With this array to collect the results we can simulate, by approximating the no. events by a Poisson variate with mean equal to the cumulative rate:

```
> for( ir in dimnames(prpw)[["RR"]] )
+ for( ip in dimnames(prpw)[["pr"]] )
+ for( iy in dimnames(prpw)[["PY"]] )
+ for( ii in dimnames(prpw)[["inc"]] )
+ {
+ nr <- as.numeric(ir)
+ np <- as.numeric(ip)/100
+ ny <- as.numeric(iy)
+ ni <- as.numeric(ii)
+ # Cumulative incidence in exposed and non-exposed:
+ cuminc <- ny * c(np,1-np) * ni * c(nr,1)
+ # Events in exposed:
+ nx <- rpois( nsim, cuminc[1] )
+ # Events in unexposed:
+ nr <- rpois( nsim, cuminc[2] )
+ # Rate Ratio
+ rr <-  (nx/np) / (nr/(1-np))
+ # Error factor
+ erf <- exp( 1.96*sqrt(1/nx+1/nr) )
+ # Significance
+ sgn <- abs(log(rr)) > log(erf)
+ # Store results
+ prpw[ir,ip,iy,ii,"RR" ] <- median( rr )
+ prpw[ir,ip,iy,ii,"Erf"] <- median( erf )
+ prpw[ir,ip,iy,ii,"Pwr"] <- mean( sgn )
+ }
```

For the sake of verification, we print the median RRs from the simulations:

```
> round( ftable( prpw[,,"12",,"RR" ],
+                 col.vars=1, row.vars=c(3,2) ), 2 )
       RR 1.05  1.1 1.15  1.2 1.25  1.3 1.35  1.4 1.45  1.5
inc pr
20  5     1.04 1.09 1.13 1.19 1.23 1.29 1.34 1.39 1.44 1.49
    10    1.05 1.09 1.14 1.20 1.24 1.30 1.34 1.40 1.44 1.49
    15    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    20    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    25    1.05 1.10 1.15 1.19 1.24 1.30 1.35 1.40 1.45 1.50
    30    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    35    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    40    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    45    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    50    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
30  5     1.04 1.09 1.14 1.20 1.25 1.28 1.35 1.39 1.44 1.49
    10    1.05 1.10 1.15 1.20 1.24 1.29 1.35 1.39 1.45 1.50
    15    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.44 1.50
    20    1.05 1.10 1.15 1.20 1.25 1.29 1.35 1.40 1.45 1.50
    25    1.05 1.10 1.15 1.20 1.25 1.30 1.34 1.40 1.45 1.50
    30    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    35    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    40    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    45    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
    50    1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50
```

Further more we see that the precision of the estimates (the error-factor) decreases only little with increasing RR, but much more so by increasing prevalence of the risk factor:

```
> round( ftable( prpw[,,c("8","12"),,"Erf"],
+                 col.vars=1, row.vars=c(3,4,2) ), 2 )
          RR 1.05  1.1 1.15  1.2 1.25  1.3 1.35  1.4 1.45  1.5
PY inc pr
8  20  5     2.03 1.96 1.96 1.96 1.90 1.89 1.85 1.84 1.84 1.80
       10    1.66 1.65 1.63 1.61 1.60 1.59 1.57 1.56 1.55 1.54
       15    1.53 1.52 1.51 1.50 1.49 1.48 1.47 1.46 1.45 1.45
       20    1.46 1.45 1.45 1.44 1.43 1.42 1.41 1.41 1.40 1.40
       25    1.42 1.41 1.41 1.40 1.39 1.39 1.38 1.37 1.37 1.36
       30    1.40 1.39 1.38 1.38 1.37 1.36 1.36 1.35 1.35 1.35
       35    1.38 1.37 1.37 1.36 1.36 1.35 1.35 1.34 1.34 1.33
       40    1.37 1.36 1.36 1.35 1.35 1.34 1.34 1.33 1.33 1.33
       45    1.36 1.36 1.35 1.35 1.34 1.34 1.34 1.33 1.33 1.33
       50    1.36 1.35 1.35 1.35 1.34 1.34 1.34 1.33 1.33 1.33
   30  5     1.79 1.75 1.72 1.71 1.69 1.68 1.66 1.64 1.64 1.62
       10    1.51 1.50 1.49 1.48 1.47 1.46 1.45 1.44 1.43 1.42
       15    1.42 1.41 1.40 1.39 1.38 1.38 1.37 1.36 1.36 1.35
       20    1.36 1.36 1.35 1.34 1.34 1.33 1.33 1.32 1.32 1.31
       25    1.33 1.33 1.32 1.32 1.31 1.31 1.30 1.30 1.29 1.29
       30    1.31 1.31 1.30 1.30 1.29 1.29 1.28 1.28 1.28 1.27
       35    1.30 1.29 1.29 1.29 1.28 1.28 1.27 1.27 1.27 1.27
       40    1.29 1.29 1.28 1.28 1.27 1.27 1.27 1.27 1.26 1.26
       45    1.29 1.28 1.28 1.27 1.27 1.27 1.27 1.26 1.26 1.26
       50    1.28 1.28 1.28 1.27 1.27 1.27 1.27 1.26 1.26 1.26
12 20  5     1.78 1.75 1.72 1.71 1.69 1.68 1.66 1.64 1.64 1.62
       10    1.51 1.50 1.49 1.48 1.47 1.46 1.45 1.44 1.43 1.42
       15    1.42 1.41 1.40 1.39 1.38 1.37 1.37 1.36 1.36 1.35
       20    1.36 1.36 1.35 1.34 1.34 1.33 1.33 1.32 1.32 1.31
       25    1.33 1.33 1.32 1.32 1.31 1.30 1.30 1.30 1.29 1.29
       30    1.31 1.31 1.30 1.30 1.29 1.29 1.28 1.28 1.28 1.27
       35    1.30 1.29 1.29 1.29 1.28 1.28 1.27 1.27 1.27 1.27
       40    1.29 1.29 1.28 1.28 1.27 1.27 1.27 1.27 1.26 1.26
       45    1.29 1.28 1.28 1.28 1.27 1.27 1.27 1.26 1.26 1.26
       50    1.28 1.28 1.28 1.28 1.27 1.27 1.27 1.26 1.26 1.26
```

```
   30  5     1.59 1.57 1.56 1.54 1.54 1.52 1.51 1.50 1.49 1.48
       10    1.40 1.39 1.38 1.37 1.37 1.36 1.35 1.35 1.34 1.33
       15    1.33 1.32 1.31 1.31 1.30 1.30 1.29 1.29 1.28 1.28
       20    1.29 1.28 1.28 1.27 1.27 1.26 1.26 1.26 1.25 1.25
       25    1.26 1.26 1.25 1.25 1.25 1.24 1.24 1.24 1.23 1.23
       30    1.25 1.24 1.24 1.24 1.23 1.23 1.23 1.22 1.22 1.22
       35    1.24 1.23 1.23 1.23 1.22 1.22 1.22 1.22 1.21 1.21
       40    1.23 1.23 1.22 1.22 1.22 1.22 1.21 1.21 1.21 1.21
       45    1.23 1.22 1.22 1.22 1.22 1.21 1.21 1.21 1.21 1.21
       50    1.23 1.22 1.22 1.22 1.22 1.21 1.21 1.21 1.21 1.21
```

We can then print the anticipated power to detect a significant RR of a given size, in percent:

```
> round( ftable( prpw[,,c("8","12"),,"Pwr"]*100,
+                col.vars=1, row.vars=c(3,4,2) ), 1 )
             RR 1.05  1.1 1.15  1.2 1.25  1.3 1.35  1.4 1.45  1.5
PY inc pr
8  20  5         4.8  6.0  7.8  9.7 11.8 15.1 18.4 21.4 25.8 28.7
       10        5.9  7.3 10.2 12.9 17.0 21.3 28.0 33.1 39.6 44.5
       15        5.9  7.9 11.5 15.8 22.1 28.2 34.9 41.3 48.9 56.3
       20        6.1  8.2 12.3 18.0 24.9 32.6 40.1 49.0 56.6 64.2
       25        6.1  8.8 13.1 19.4 26.9 35.7 44.8 53.7 62.5 70.8
       30        6.3  8.8 13.6 20.5 28.4 37.3 47.8 58.8 67.0 75.5
       35        6.3  9.7 15.3 21.4 30.3 40.6 50.4 60.9 69.8 77.6
       40        6.2  9.3 14.9 22.5 31.7 41.6 53.3 63.6 71.8 80.2
       45        6.2  9.3 15.3 22.0 31.3 42.0 53.8 63.6 73.1 81.1
       50        5.9  9.4 15.2 22.2 31.3 42.4 53.4 63.9 72.9 81.3
   30  5         5.5  7.3  9.6 11.8 15.2 19.8 22.8 28.1 33.7 37.5
       10        6.2  8.4 11.6 16.0 22.6 29.0 36.8 43.5 51.4 58.7
       15        6.3  8.9 13.7 20.7 28.5 36.7 46.6 55.1 65.0 72.5
       20        6.7 10.4 15.8 24.1 33.1 43.7 54.2 63.4 73.2 81.0
       25        6.8 10.4 17.1 26.4 37.9 47.9 60.9 70.1 79.9 86.5
       30        7.1 11.3 19.0 28.8 40.7 52.1 64.4 75.0 83.2 89.7
       35        6.7 11.1 19.1 30.2 42.2 54.8 67.6 78.3 86.2 91.7
       40        6.9 11.3 19.8 30.8 44.4 58.0 70.1 80.3 87.5 92.9
       45        7.1 11.3 20.1 31.0 45.3 58.1 70.7 81.3 87.9 93.4
       50        6.5 11.3 19.9 31.1 44.3 57.4 70.5 81.3 88.6 93.4
12 20  5         5.3  6.9  9.1 12.1 14.3 20.0 23.2 28.8 33.5 38.9
       10        5.8  8.3 11.7 17.4 22.8 29.2 35.7 44.1 52.0 59.0
       15        6.2  9.5 13.8 20.4 28.4 37.5 46.9 56.2 64.8 72.4
       20        7.0 10.5 16.0 23.4 33.1 43.7 54.1 64.7 73.1 81.7
       25        7.0 10.5 17.4 25.0 36.6 48.8 60.8 70.9 78.8 85.7
       30        6.3 12.0 18.6 27.9 41.2 52.3 64.5 75.4 83.2 89.8
       35        6.9 11.5 19.4 30.0 41.8 55.5 67.5 78.7 86.1 91.5
       40        6.9 11.8 19.9 30.8 43.4 57.6 70.3 80.3 87.6 92.7
       45        6.4 11.6 20.1 30.7 44.4 57.8 70.2 81.2 88.5 93.6
       50        6.7 11.5 20.3 31.7 45.3 58.2 70.6 80.3 89.0 93.5
   30  5         6.0  8.1 10.6 15.0 19.3 25.3 31.6 37.5 44.3 51.0
       10        6.6  9.8 15.2 21.6 30.3 38.9 49.6 58.5 67.6 74.9
       15        6.8 11.2 18.1 27.4 39.0 49.9 61.2 72.1 79.9 86.9
       20        6.8 11.8 20.6 32.0 45.0 57.8 71.4 80.7 88.1 93.0
       25        7.1 13.6 23.7 35.7 51.0 64.6 76.2 86.0 92.3 96.3
       30        7.0 14.4 24.7 39.1 55.1 69.2 81.5 89.3 94.5 97.5
       35        7.5 14.4 26.5 41.4 58.2 71.8 83.1 91.2 96.1 98.2
       40        7.0 14.4 27.2 43.4 59.8 74.2 85.3 92.3 96.5 98.6
       45        7.6 15.2 27.7 44.2 60.3 75.5 85.8 93.2 97.1 99.0
       50        7.9 15.0 27.5 44.3 61.7 75.3 86.9 93.3 97.0 99.0
```

And show where the power exceeds 80%:

```
> print( ftable( (prpw[,,c("8","12"),,"Pwr"]>0.8)*1,
+                col.vars=1, row.vars=c(3,4,2) ),
+        zero.print="." )
```

```
          RR 1.05 1.1 1.15 1.2 1.25 1.3 1.35 1.4 1.45 1.5
PY inc pr
8  20  5        .   .    .   .    .   .    .   .    .   .
       10       .   .    .   .    .   .    .   .    .   .
       15       .   .    .   .    .   .    .   .    .   .
       20       .   .    .   .    .   .    .   .    .   .
       25       .   .    .   .    .   .    .   .    .   .
       30       .   .    .   .    .   .    .   .    .   .
       35       .   .    .   .    .   .    .   .    .   .
       40       .   .    .   .    .   .    .   .    .   1
       45       .   .    .   .    .   .    .   .    .   1
       50       .   .    .   .    .   .    .   .    .   1
   30  5        .   .    .   .    .   .    .   .    .   .
       10       .   .    .   .    .   .    .   .    .   .
       15       .   .    .   .    .   .    .   .    .   .
       20       .   .    .   .    .   .    .   .    .   1
       25       .   .    .   .    .   .    .   .    .   1
       30       .   .    .   .    .   .    .   .    1   1
       35       .   .    .   .    .   .    .   .    1   1
       40       .   .    .   .    .   .    .   1    1   1
       45       .   .    .   .    .   .    .   1    1   1
       50       .   .    .   .    .   .    .   1    1   1
12 20  5        .   .    .   .    .   .    .   .    .   .
       10       .   .    .   .    .   .    .   .    .   .
       15       .   .    .   .    .   .    .   .    .   .
       20       .   .    .   .    .   .    .   .    .   1
       25       .   .    .   .    .   .    .   .    .   1
       30       .   .    .   .    .   .    .   .    1   1
       35       .   .    .   .    .   .    .   .    1   1
       40       .   .    .   .    .   .    .   1    1   1
       45       .   .    .   .    .   .    .   1    1   1
       50       .   .    .   .    .   .    .   1    1   1
   30  5        .   .    .   .    .   .    .   .    .   .
       10       .   .    .   .    .   .    .   .    .   .
       15       .   .    .   .    .   .    .   .    .   1
       20       .   .    .   .    .   .    .   1    1   1
       25       .   .    .   .    .   .    .   1    1   1
       30       .   .    .   .    .   .    1   1    1   1
       35       .   .    .   .    .   .    1   1    1   1
       40       .   .    .   .    .   .    1   1    1   1
       45       .   .    .   .    .   .    1   1    1   1
       50       .   .    .   .    .   .    1   1    1   1
```

Broadly speaking, with 80% power we can only see effects of RR> 1.4 for exposures with prevalence over 30%, but the good news is that with 12,000 PY we have a relative precision of 1.3 and with 8,000 PY a relative precision in the vicinity of 1.5.