A first introduction to on ADRminer

Ismaïl Ahmed

23 juillet 2014

This tutorial aims at illustrate the functionalities of the ADRminer package. It is based on spontaneous reports from the FAERS. The data that will be used throughout the document are from the second and third semester of 2013 and can be downloaded [here][id] (the ASCII version). We assume that the data are uncompressed and gathered in a common directory

```
list.files()
```

```
## [1] "aers_ascii_2012q2"  "AERS_ASCII_2013q3"  "faers_ascii_2012q4"
## [4] "faers_ascii_2013q1" "faers_ascii_2013q2" "Tutorial.knit.md"
```

Data import

The first step is to import the data and to convert them into objects that can be handled by the ADRminer package. ADRminer provides with a function which facilitate the importation of the data from the AERS This function requires at least three arguments :

1. a path to the drug file
2. a path to the adverse reaction file
3. a path to the demo file, the latter being used to collect individuals characteristics such as the age and gender of the patients

```
drugFile <- c("./faers_ascii_2013q2/asii/DRUG13Q2.txt")
reacFile <- c("./faers_ascii_2013q2/asii/REAC13Q2.txt")
demoFile <- c("./faers_ascii_2013q2/asii/DEMO13Q2.txt")

library(ADRminer)
```

```
## Loading required package: Matrix
## Loading required package: glmnet
## Loaded glmnet 1.9-8
##
## Loading required package: data.table
## Loaded ADRminer 0.80
```

```
faers13q2 <- readFAERS(drugFile, reacFile, demoFile)
faers13q2
```

```
##
##  S4 class: pvInd
## @drug: Sparse matrix:  53356 x 8549
## @ae: Sparse matrix 53356 x 6286
## @cov: Covariate data.frame: 53356 x 5
## primaryid age gndr_cod reporter_country occp_cod
```

Using the default value of readFAERS, we see that the `faers13q2` is an object of class S4 `pvInd` , which is made of

1. a `sparse matrix` containing 53356 spontaneous reports and involvinv x 8549 different drugs
2. a `sparse matrix` containing 53356 spontaneous reports and involving x 8549 different adverse events
3. a `data.frame` containing indivdual covariates extracted from the DEMO**.txt file

it is really unlikely that a signal generation procedure will be restricted on one semester. Accordingly we extended the `rbind` function to the class pvInd in order to merge several pvInd object

```
## data import from the third semester 2013
drugFile <- c("./AERS_ASCII_2013q3/ASCII/DRUG13Q3.txt")
reacFile <- c("./AERS_ASCII_2013q3/ASCII/REAC13Q3.txt")
demoFile <- c("./AERS_ASCII_2013q3/ASCII/DEMO13Q3.txt")
faers13q3 <- readFAERS(drugFile, reacFile, demoFile)
faers13q3
```

```
##
##  S4 class: pvInd
## @drug: Sparse matrix:  52157 x 8955
## @ae: Sparse matrix 52157 x 6322
## @cov: Covariate data.frame: 52157 x 5
## primaryid age gndr_cod reporter_country occp_cod
```

```
faers13q23 <- rbind(faers13q2, faers13q3)
```

```
##
##  S4 class: pvInd
## @drug: Sparse matrix:  53356 x 8549
## @ae: Sparse matrix 53356 x 6286
## @cov: Covariate data.frame: 53356 x 5
## primaryid age gndr_cod reporter_country occp_cod
##
##  S4 class: pvInd
## @drug: Sparse matrix:  52157 x 8955
## @ae: Sparse matrix 52157 x 6322
## @cov: Covariate data.frame: 52157 x 5
## primaryid age gndr_cod reporter_country occp_cod
```

```
faers13q23
```

```
##
##  S4 class: pvInd
## @drug: Sparse matrix:  105512 x 13036
## @ae: Sparse matrix 105512 x 7930
## @cov: Covariate data.frame: 105512 x 5
## primaryid age gndr_cod reporter_country occp_cod
```

We see that spontaneous reports are concatenated. Note that the number of drugs and adverse events increase as both file do not involve the same drugs and adverse events.

Some basic functions have been developed to access and manipulate the object `faers13q23`. In particular, we can use the `pvIndResize` function to eliminate drugs and adverse events associated with a too small (say less than 50) number of spontaneous reports which will drastically decrease the number of drugs and adverse events.

```
faers13q23resize <- pvIndResize(faers13q23, aeMarginMin = 50, drugMarginMin = 50)
faers13q23resize
```

```
##
##  S4 class: pvInd
## @drug: Sparse matrix:  105512 x 543
## @ae: Sparse matrix 105512 x 860
## @cov: Covariate data.frame: 105512 x 5
## primaryid age gndr_cod reporter_country occp_cod
```

Signal detection analysis

Gamma Poisson Shrinker

The first method illustrated in this tutorial is the Gamma Poisson Shrinker (GPS) initially proposed by DuMouchel (American Statistician 1999) as well as its extension to the multiple comparison framework (Ahmed et al. Stat Med 2009). The corresponding function is `gps` which can take a number of arguments. The default parametrisation corresponds to the extension proposed in Ahmed et al. (2009): the drug adverse event pairs are ranked according to the posterior probability of a null hypothesis (H0: the relatvie risk `rr0 <= 1` and a signal is generated based on an estimated False Discovery Rate (FDR, Benjamini an Hochberg JRSSB 1995) less than 0.05. Alternatively, it is also possible to use the detection strategy proposed in Szarfman et al. (Drug Safety 2002) which consists in highlighting drug-ae pairs associated with EB05 > 2.

```
resGPSpH0 <- gps(faers13q23resize)
## equivalent to resGPSpH0 <- gps(faers13q23resize, assocMeasure = "posH0",
## detectCriter = "FDR", criterThres = 0.05))
```

The number of generated signals is:

```
resGPSpH0$nSig
```

```
## [1] 12581
```

A summary of the characteristics of the generated signals is stored in `$sig`

```
head(resGPSpH0$sig)
```

```
##                drug                      event   n expected postH0    rrr
## 1        ALPRAZOLAM        Accidental overdose 184   1.0405      0 176.84
## 2           Avandia Cardiac failure congestive 355   7.6285      0  46.54
## 3           Avandia      Myocardial infarction 525  15.0176      0  34.96
## 4 CLOZAPINE TABLETS          Granulocytopenia 290   1.0649      0 272.32
## 5          Diazepam        Exposure via father 185   0.4113      0 449.84
## 6          Diazepam       Small for dates baby 178   0.4074      0 436.92
##   drugMargin aeMargin FDR postH0
```

```
## 1        555       435    0       0
## 2       2525       701    0       0
## 3       2525      1380    0       0
## 4        765       323    0       0
## 5        448       213    0       0
## 6        448       211    0       0
```

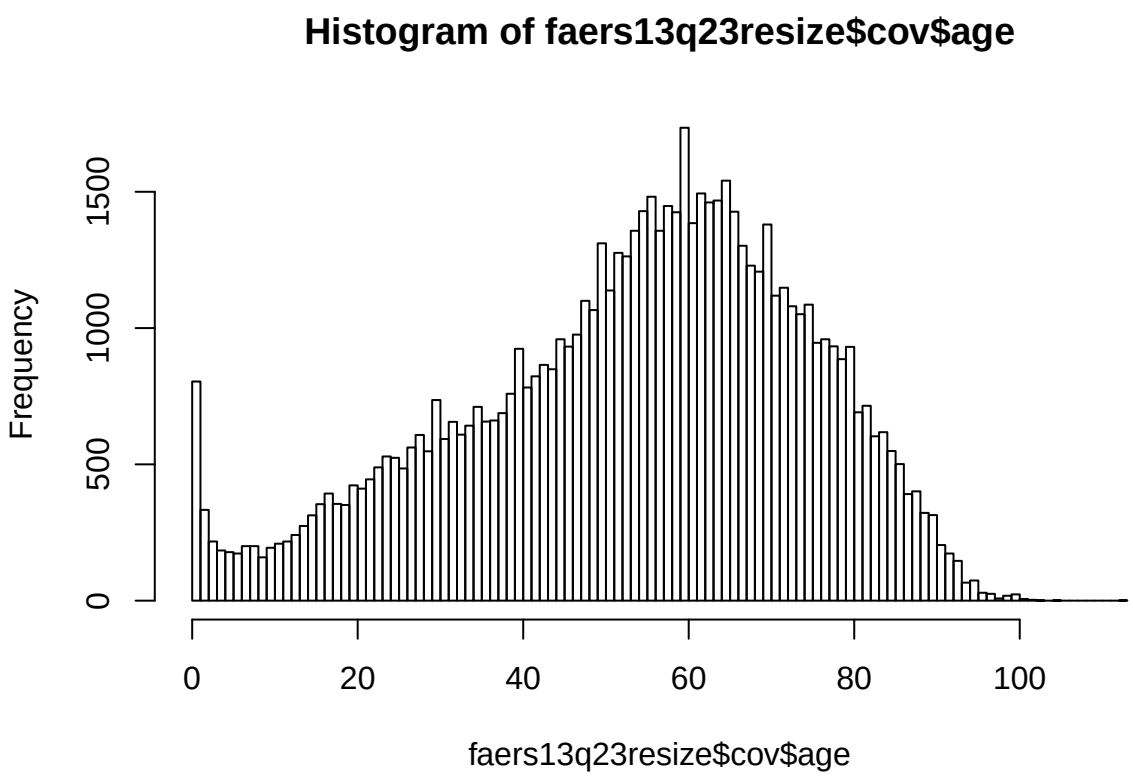
Here are the results obtained with the detection strategy proposed by Szarfman et al.

```
resGPSsz <- gps(faers13q23resize, assocMeasure = "lb05",
                detectCriter = "assocMeasure", criterThres = 2)
resGPSsz$nSig
head(resGPSsz$sig)
```

It is also possible ta perform stratified GPS analysis according to covariates stored in the pvInd object "faers13q23resize$cov". However, this may require to recode some covariates into `factors` with a reasonable number of categories especially if one is willing to stratify according to several covariates such as age and gender.

Here we illustrate how to recode the age covariate into a new factor variable

```
hist(faers13q23resize$cov$age, breaks = 100)
```




```
ageFac <- cut(faers13q23resize$cov$age, c(0,1,10,25,50, 120), include.lowest = T)
table(ageFac)
```

```
## ageFac
##    [0,1]   (1,10]  (10,25]  (25,50] (50,120]
##      804     1838     5528    19502    43826
```

```
ageFac <- addNA(ageFac) ## this is to consider NA values as a category
table(ageFac)
```

```
## ageFac
##    [0,1]   (1,10]  (10,25]  (25,50] (50,120]     <NA>
##      804     1838     5528    19502    43826    34014
```

```
faers13q23resize$cov$ageFac <- ageFac
```

The syntax to run stratified GPS according to ageFac is then

```
resGPSstrat <- gps(faers13q23resize, strat="ageFac")
```

```
## varStrat
##   (1,10]  (10,25]  (25,50] (50,120]    [0,1]       NA
##     1838     5528    19502    43826      804    34014
```

```
resGPSstrat$nSig
```

```
## [1] 12512
```

```
head(resGPSstrat$sig)
```

```
##                drug                              event    n expected postH0
## 1           Avandia        Cardiac failure congestive  355    9.334      0
## 2           Avandia             Myocardial infarction  525   17.074      0
## 3 CLOZAPINE TABLETS                  Granulocytopenia  290    1.505      0
## 4            Enbrel           Injection site erythema 1186  180.990      0
## 5            Enbrel              Injection site pain 1402  215.582      0
## 6         Methadose Drug withdrawal syndrome neonatal  251    3.210      0
##       rrr drugMargin aeMargin FDR postH0
## 1  38.034       2525      701   0      0
## 2  30.748       2525     1380   0      0
## 3 192.684        765      323   0      0
## 4   6.553      22085     1447   0      0
## 5   6.503      22085     1762   0      0
## 6  78.205        524      435   0      0
```