Package 'PhViD'

February 19, 2015

Type Package

Title PhViD: an R package for PharmacoVigilance signal Detection.

Version 1.0.6

Date 2013-04-30

Author Ismaïl Ahmed & Antoine Poncet

Maintainer Ismaïl Ahmed <ismail.ahmed@inserm.fr>

Depends R (>= 2.10), LBE, MCMCpack

Description PhViD contains several pharmacovigilance signal detection methods extended to the multiple comparison setting.

License GPL-2

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2013-04-30 18:52:52

SystemRequirements Tcl/Tk package TkTable.

R topics documented:

PhViD-package *PhViD: a R package for PharmacoVigilance signal Detection*

Description

The PhViD-package proposes the main pharmacovigilance signal detection methods extended to the multiple comparison setting. For the frequentist methods, the package requires the LBE procedure that is stored in the Bioconductor website http://bioconductor.org/. LBE can be installed by entering

```
source("http://bioconductor.org/biocLite.R")
biocLite("LBE")
```

in the R console.

Author(s)

Ismaïl Ahmed & Antoine Poncet

Maintainer: Ismaïl Ahmed <ismail.ahmed@inserm.fr>

References

Ahmed I, Thiessard F, Miremont-Salamé G, Bégaud B, Tubert-Bitter P. Pharmacovigilance data mining with methods based on false discovery rates: a comparative simulation study. Clin. Pharmacol. Ther. 2010 Oct;88(4):492-498.

Ahmed I, Dalmasso C, Haramburu F, Thiessard F, Broët P, Tubert-Bitter P. False discovery rate estimation for frequentist pharmacovigilance signal detection methods. Biometrics. 2010 Mar;66(1):301-309.

Ahmed I, Haramburu F, Fourrier-Réglat A, Thiessard F, Kreft-Jais C, Miremont-Salamé G, Bégaud B, Tubert-Bitter P. Bayesian pharmacovigilance signal detection methods revisited in a multiple comparison setting. Stat Med. 2009 Jun 15;28(13):1774-1792.

Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, De Freitas RM. A Bayesian Neural Network Method for Adverse Drug Reaction Signal Generation European Journal of Clinical Pharmacology, 1998, 54, 315-321.

Dalmasso C, Broet P, Moreau T (2005), A simple procedure for estimating the false discovery rate, Bioinformatics, Bioinformatics, 21: 660 - 668.

DuMouchel W. Bayesian Data Mining in Large Frequency Tables, with an Application to the FDA Spontaneous Reporting System. The American Statistician. 1999, 53. 177-190.

Evans SJ, Waller PC, Davis S, Use of Proportional Reporting Ratios (PRRs) for Signal Generation from Spontaneous Adverse Drug Reaction Reports Pharmacoepidemiology and Drug Safety, 2001, 10, 483-486.

Noren, GN, Bate A, Orre R, Edwards IR, Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events Statistics in Medicine, 2006, 25, 3740-3757.

van Puijenbroek EP, Bate A, Leufkens HG, Lindquist M, Orre R, Egberts AC, A Comparison of Measures of Disproportionality for Signal Detection in Spontaneous Reporting Systems for Adverse Drug Reactions Pharmacoepidemiology and Drug Safety, 2002, 11, 3-1.

as.PhViD *data.frame to PhViD data*

Description

as.PhViD is a function that converts a data.frame into an object that can be used in the signal detection method functions.

Usage

```
as.PhViD(DATA.FRAME, MARGIN.THRES = 1)
```

Arguments

DATA.FRAME	The data.frame has to be structured as follows 1st column: label of the drugs 2nd column: label of the adverse events 3rd column: Number of spontaneous reports of the corresponding couple drug-adverse event.
MARGIN.THRES	This option can be used to eliminate the drugs and the adverse events for which the marginal counts are less than MARGIN.THRES.

Value

L	data.frame that contains the labels of the drugs and the adverse events.
N	sum of the spontaneous reports counts.
data	data.frame that contains the number of spontaneous reports (n11) and the corresponding marginal counts as well (n1. and n.1).

Author(s)

Ismaïl Ahmed & Antoine Poncet

BCPNN *Bayesian confidence propagation neural network*

Description

Bayesian confidence propagation neural network (Bate et al. 1998, Noren et al. 2006) extended to the multiple comparison framework.

Usage

```
BCPNN(DATABASE, RR0 = 1, MIN.n11 = 1, DECISION = 1, DECISION.THRES = 0.05,
RANKSTAT = 1, MC = FALSE, NB.MC = 10000)
```

Arguments

DATABASE	Object returned by the function as.PhViD.
RR0	Value of the tested risk. By default, RR0=1.
MIN.n11	Minimum number of notifications for a couple to be potentially considered as a signal. By default, MIN.n11 = 1.
DECISION	Decision rule for the signal generation based on 1 = FDR (Default value) 2 = Number of signals 3 = Ranking statistic. See RANKSTAT
DECISION.THRES	Threshold for DECISION. Ex 0.05 for FDR (DECISION=1).
RANKSTAT	Statistic used for ranking the couples: 1 = Posterior probability of the null hypothesis 2 = 2.5% quantile of the posterior distribution of IC.
MC	If MC=TRUE, the statistic of interest (see RANKSTAT) is calculated by Monte Carlo simulations which can be very long. If MC=FALSE, IC is approximated by a normal distribution (which can be very crude for small counts).
NB.MC	If MC=TRUE, NB.MC indicates the number of Monte Carlo simulations to be done

Details

The BCPNN method is based on the calculation of the Information Component IC. If MC = FALSE, the bayesian model used is the beta-binomial proposed by Bate et al. (1998). The statistic of interest (see RANKSTAT) is calculated by the normal approximation made in Bate et al. (1998) with the use of the exact expectation and variance proposed by Gould (2003). If MC = TRUE, the model is based on the Dirichlet-multinomial model proposed more recently in Noren et al. (2006). In this case, the statistic of interest is calculated by Monte Carlo simulations.

Value

ALLSIGNALS	Data.frame summarizing the results of all couples with at least MIN.n11 notifications ordered by RANKSTAT. It contains notably the labels, the cell counts, the expected counts ($n1.\times n.1/N$, see as.PhViD), RANKSTAT, the ratios(count/expected count), the marginal counts and the estimations of FDR, FNR, Se et Sp. If RANKSTAT!=1, the last column is the posterior probability of the null hypothesis.
SIGNALS	Same Data.frame as ALLSIGNALS but restricted to the list of generated signals.
NB.SIGNALS	Number of generated signals.
INPUT.PARAM	Parameters entered in the function.

Author(s)

Ismaïl Ahmed & Antoine Poncet

References

Ahmed I, Haramburu F, Fourrier-Réglat A, Thiessard F, Kreft-Jais C, Miremont-Salamé G, Bégaud B, Tubert-Bitter P. Bayesian pharmacovigilance signal detection methods revisited in a multiple comparison setting. Stat Med. 2009 Jun 15;28(13):1774-1792.

Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, De Freitas RM, A Bayesian Neural Network Method for Adverse Drug Reaction Signal Generation European Journal of Clinical Pharmacology, 1998, 54, 315-321.

Gould AL, Practical Pharmacovigilance Analysis Strategies Pharmacoepidemiology and Drug Safety, 2003, 12, 559-574

Noren, GN, Bate A, Orre R, Edwards IR, Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events Statistics in Medicine, 2006, 25, 3740-3757.

Examples

```
## start
data(PhViDdata.frame)
PhViDdata <- as.PhViD(PhViDdata.frame)
# res <- BCPNN(PhViDdata)
## end
```

GPS *Gamma Poisson Shrinkage*

Description

Gamma Poisson Shrinkage model proposed by DuMouchel (1999) extended to the multiple comparison framework.

Usage

```
GPS(DATABASE,  RR0 = 1, MIN.n11 = 1, DECISION = 1, DECISION.THRES = 0.05,
RANKSTAT = 1, TRONC = FALSE, TRONC.THRES = 1,
PRIOR.INIT = c(alpha1 = 0.2, beta1 = 0.06, alpha2 = 1.4,
beta2 = 1.8, w = 0.1), PRIOR.PARAM = NULL)
```

Arguments

DATABASE	Object returned by the function as.PhViD.
RR0	Value of the tested risk. By default, RR0=1.
MIN.n11	Minimum number of notifications for a couple to be potentially considered as a signal. This option does not affect the calculation of the hyper parameters. By default, MIN.n11 = 1.

DECISION	Decision rule for the signal generation based on 1 = FDR (Default value) 2 = Number of signals 3 = Ranking statistic. See RANKSTAT
DECISION.THRES	Threshold for DECISION. Ex 0.05 for FDR (DECISION=1).
RANKSTAT	Statistic used for ranking the couples: 1 = Posterior probability of the null hypothesis 2 = 5% quantile of the posterior distribution of λ 3 = Posterior Expectation of $\log_2(\lambda)$
TRONC	If TRUE, only the data with at least TRONC.THRES notifications are considered in the calculation of the hyper parameters and the likelihood is a product of mixture of two negative binomial truncated by TRONC.THRES-1. By default, TRONC=F
TRONC.THRES	See TRONC
PRIOR.INIT	Vector of initialization of the prior parameters $(\alpha_1, \beta_1, \alpha_2, \beta_2, w)$. By default, PRIOR.INIT $= c(\alpha_1 = 0.2, \beta_1 = 0.06, \alpha_2 = 1.4, \beta_2 = 1.8, w = 0.1)$, ie the prior parameters found in DuMouchel (1999).
PRIOR.PARAM	Chosen hyper parameters. By default, PRIOR.PARAM = NULL which means that the hyperparameters are calculated by maximising the marginal likelihood.

Details

Each observed count n_{11} is assumed to be drawn from a Poisson distribution with parameters e_{11} where e_{11} is the expected count under the hypothesis of independence between the adverse events and the drugs ($n1. \times n.1/N$, see as.PhViD). λ is a priori assumed to be distributed according to a mixture of two gamma distributions: $\lambda \sim w\,\Gamma(\alpha_1, \beta_1) + (1-w)\,\Gamma(\alpha_2, \beta_2)$.

Value

ALLSIGNALS	Data.frame summarizing the results of all couples with at least MIN.n11 notifications ordered by RANKSTAT. It contains notably the labels, the cell counts, the expected counts, RANKSTAT, the ratios(count/expected count), the marginal counts and the estimations of FDR, FNR, Se et Sp. If RANKSTAT!=1, the last column is the posterior probability of the null hypothesis.
SIGNALS	Same Data.frame as ALLSIGNALS but restricted to the list of generated signals.
NB.SIGNALS	Number of generated signals.
INPUT.PARAM	Parameters entered in the function.
PARAM	A list that contains the prior hyper parameters (PRIOR.PARAM). Additionally if PRIOR.PARAM=NULL, it also contains the prior hyper parameters initialization (PRIOR.INIT) and the convergence code (see nlm()).

Author(s)

Ismaïl Ahmed & Antoine Poncet

References

Ahmed I, Haramburu F, Fourrier-Réglat A, Thiessard F, Kreft-Jais C, Miremont-Salamé G, Bégaud B, Tubert-Bitter P. Bayesian pharmacovigilance signal detection methods revisited in a multiple comparison setting. Stat Med. 2009 Jun 15;28(13):1774-1792.

DuMouchel W, Bayesian Data Mining in Large Frequency Tables, with an Application to the FDA Spontaneous Reporting System, The American Statistician, 1999, 53, 177-190.

Szarfman A, Machado S, O'Neill R, Use of Screening Algorithms and Computer Systems to Efficiently Signal Higher-Than-Expected Combinations of Drugs and Events in the US FDA's Spontaneous Reports Database Drug Safety, 2002, 25, 381-392.

Examples

```
## start
#data(PhViDdata.frame)


#PhViDdata <- as.PhViD(PhViDdata.frame)
#res <- GPS(PhViDdata)

#List of signals generated by the decision rule proposed
#by Szarfman et al. (2002)
#res2 <- GPS(PhViDdata, DECISION = 3, DECISION.THRES = 2, RANKSTAT = 2)
## end
```

`PhViD.search` *PhViD.search*

Description

This function makes possible to extract some information from the output of the PhViD functions for a given couple adverse event-drug, for a drug or for an adverse event.

Usage

```
PhViD.search(RESULT, DRUG = NULL, EVENT = NULL)
```

Arguments

`RESULT`	`RESULT` must be the output of one the signal detection method functions (`ROR, PRR, RFET, GPS` or `BCPNN`
`DRUG`	The label of the drug. By default, `DRUG=FALSE`.
`EVENT`	The label of the adverse event. By default, `EVENT=FALSE`.

Value

DRUG	Recalls the label of the drug.
EVENT	Recalls the label of the event.
EXIST_DRUG	Indicates if the label of the drug exists in the database.
EVENT	Indicates if the label of the adverse event exists in the database.
EXIST_COUPLE	Indicates if the couple is present in the database.
LIST	It is a dataframe that contains the labels, the counts, the expected counts, the value of the statistic of interest, the rank and the estimated FDR for each couple.

Author(s)

Antoine Poncet & Ismaïl Ahmed

PhViDdata.frame *Simulated Pharmacovigilance data*

Description

This is a simulated data set aiming at mimicking the French database coded in ATC5 for the drugs and HLT for the adverse events. The simulation procedure is described in Ahmed et al.

Usage

```
data(PhViDdata.frame)
```

Format

A data frame with 102483 observations on the following 3 variables.

`Drug lab` a factor indicating the label of the 634 drugs.

`AE lab` a factor indicating the label of the 756 adverse events.

`n11` a numeric vector indicating the number of spontaneous reports of the corresponding couple.

Author(s)

Ismaïl Ahmed & Antoine Poncet

PRR *Proportional Reporting Ratio*

Description

Proportional Reporting Ratio proposed by Evans et al. (2001) extended to the multiple comparison framework. Note that the computed variance is different from the one used in van Puijenbroek et al. (2002)

Usage

```
PRR(DATABASE, RR0 = 1, MIN.n11 = 1, DECISION = 1,
DECISION.THRES = 0.05, RANKSTAT = 1)
```

Arguments

DATABASE	Object returned by the function as.PhViD.
RR0	Value of the tested relative risk. By default, RR0=1.
MIN.n11	Minimum number of notifications for a couple to be potentially considered as a signal. By default, MIN.n11 = 1.
DECISION	Decision rule for the signal generation based on 1 = FDR (Default value) 2 = Number of signals 3 = Ranking statistic. See RANKSTAT
DECISION.THRES	Threshold for DECISION. Ex 0.05 for FDR (DECISION=1).
RANKSTAT	Statistic used for ranking the couples: 1 = P-value 2 = Lower bound of the 95% two sided confidence interval of log(PRR).

Details

The FDR is estimated with the LBE procedure proposed by Dalmasso et al. (2005). Note that the FDR can only be estimated if the statistic of interest is the P-value.

Value

ALLSIGNALS	Data.frame summarizing the results of all couples with at least MIN.n11 notifications ordered by RANKSTAT. It contains notably the labels, the cell counts, the expected counts ($n1. \times n.1/N$, see as.PhViD), RANKSTAT, the observed relative risks (PRR), the marginal counts and the estimations of FDR (when RANKSTAT=1.)
SIGNALS	Same Data.frame as ALLSIGNALS but restricted to the list of generated signals.
NB.SIGNALS	Number of generated signals.
INPUT.PARAM	Parameters entered in the function.

Author(s)

Ismaïl Ahmed & Antoine Poncet

References

Ahmed I, Dalmasso C, Haramburu F, Thiessard F, Broët P, Tubert-Bitter P. False discovery rate estimation for frequentist pharmacovigilance signal detection methods. Biometrics. 2010 Mar;66(1):301-309.

Dalmasso C, Broët P, Moreau T (2005), A simple procedure for estimating the false discovery rate, Bioinformatics, Bioinformatics, 21: 660 - 668.

Evans SJ, Waller PC, Davis S, Use of Proportional Reporting Ratios (PRRs) for Signal Generation from Spontaneous Adverse Drug Reaction Reports Pharmacoepidemiology and Drug Safety, 2001, 10, 483-486.

van Puijenbroek EP, Bate A, Leufkens HGM, Lindquist M, Orre R and Egberts ACG, A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions, Pharmacoepidemiology and Drug Safety, 2002, 11, 3-10.

Examples

```
## start
data(PhViDdata.frame)
PhViDdata <- as.PhViD(PhViDdata.frame)
# res <- PRR(PhViDdata)
## end
```

RFET *Reporting Fisher's Exact Test*

Description

This function proposes the Fisher's Exact Test as an alternative to the PRR and ROR methods. The statistic of interest is the P-value or the mid-P-value resulting from the test (Ahmed et al., Biometrics).

Usage

```
RFET(DATABASE, OR0 = 1, MIN.n11 = 1, DECISION = 1,
DECISION.THRES = 0.05, MID.PVAL = FALSE)
```

Arguments

DATABASE	Object returned by the function as.PhViD.
OR0	Value of the tested odds ratio. By default, OR0=1.
MIN.n11	Minimum number of notifications for a couple to be potentially considered as a signal. By default, MIN.n11 = 1.

DECISION	Decision rule for the signal generation based on 1 = FDR (Default value) 2 = Number of signals 3 = P-values or mid-P-values. See MID.PVAL
DECISION.THRES	Threshold for DECISION. Ex 0.05 for FDR (DECISION=1).
MID.PVAL	if MID.PVAL=TRUE, the statistic of interest becomes the mid-P-values instead of the P-values resulting from the Fisher's exact test. By default MID.PVAL=FALSE.

Details

The FDR is estimated with the LBE procedure proposed by Dalmasso et al. (2005).

Value

ALLSIGNALS	Data.frame summarizing the results of all couples with at least MIN.n11 notifications ordered by RANKSTAT. It contains notably the labels, the cell counts, the expected count ($n1. \times n.1/N$, see as.PhViD), RANKSTAT, the observed odds ratio (ROR), the marginal counts and the estimation of FDR.
SIGNALS	Same Data.frame as ALLSIGNALS but restricted to the list of generated signals.
NB.SIGNALS	Number of generated signals.
INPUT.PARAM	Parameters entered in the function.

Author(s)

Ismaïl Ahmed & Antoine Poncet

References

Ahmed I, Dalmasso C, Haramburu F, Thiessard F, Broët P, Tubert-Bitter P. False discovery rate estimation for frequentist pharmacovigilance signal detection methods. Biometrics. 2010 Mar;66(1):301-309.

Dalmasso C, Broët P, Moreau T (2005), A simple procedure for estimating the false discovery rate, Bioinformatics, Bioinformatics, 21: 660 - 668.

Examples

```
## start
#data(PhViDdata.frame)
#PhViDdata <- as.PhViD(PhViDdata.frame)
#res <- RFET(PhViDdata)
## end
```

ROR *Reporting Odds Ratio*

Description

Reporting Odds Ratio proposed by van Puijenbroak et al. (2002) extended to the multiple comparison framework.

Usage

```
ROR(DATABASE, OR0 = 1, MIN.n11 = 1, DECISION = 1,
DECISION.THRES = 0.05, RANKSTAT = 1)
```

Arguments

DATABASE	Object returned by the function as.PhViD.
OR0	Value of the tested odds ratio. By default, OR0=1.
MIN.n11	Minimum number of notifications for a couple to be potentially considered as a signal. By default, MIN.n11 = 1.
DECISION	Decision rule for the signal generation based on 1 = FDR (Default value) 2 = Number of signals 3 = Ranking statistic. See RANKSTAT
DECISION.THRES	Threshold for DECISION. Ex 0.05 for FDR (DECISION=1).
RANKSTAT	Statistic used for ranking the couples: 1 = P-value 2 = Lower bound of the 95% two sided confidence interval of log(ROR).

Details

The FDR is estimated with the LBE procedure proposed by Dalmasso et al. (2005). Note that the FDR can only be estimated if the statistic of interest is the P-value.

Value

ALLSIGNALS	Data.frame summarizing the results of all couples with at least MIN.n11 notifications ordered by RANKSTAT. It contains notably the labels, the cell counts, the expected counts ($n1. \times n.1/N$, see as.PhViD), RANKSTAT, the observed odds ratios (ROR), the marginal counts and the estimations of FDR (when RANKSTAT=1.)
SIGNALS	Same Data.frame as ALLSIGNALS but restricted to the list of generated signals.
NB.SIGNALS	Number of generated signals.
INPUT.PARAM	Parameters entered in the function.

Author(s)

Ismaïl Ahmed & Antoine Poncet

References

Ahmed I, Dalmasso C, Haramburu F, Thiessard F, Broët P, Tubert-Bitter P. False discovery rate estimation for frequentist pharmacovigilance signal detection methods. Biometrics. 2010 Mar;66(1):301-309.

Dalmasso C, Broët P, Moreau T (2005), A simple procedure for estimating the false discovery rate, Bioinformatics, Bioinformatics, 21: 660 - 668.

van Puijenbroek EP, Bate A, Leufkens HG, Lindquist M, Orre R, Egberts AC, A Comparison of Measures of Disproportionality for Signal Detection in Spontaneous Reporting Systems for Adverse Drug Reactions Pharmacoepidemiology and Drug Safety, 2002, 11, 3-1.

Examples

```
## start
data(PhViDdata.frame)
PhViDdata <- as.PhViD(PhViDdata.frame)
res <- ROR(PhViDdata, MIN.n11 = 3)

# Decision rule proposed by van Puijenbroek et al. (2002)
# res2 <- ROR(PhViDdata, MIN.n11 = 1, DECISION=3, DECISION.THRES=0, RANKSTAT=2)
## end
```

Index