Package 'exposomex'

December 13, 2022

Title One integrated platform for exposomic analysis.

Version 1.0.0

Description
One integrated platform ``ExpsomeX'' for exposomic analysis (see: http://www.exposomex.cn/), including 14 sub-packages: exstat,exviz,exdb,excros,exmo,exmeta,exmedt,exnta,exsurv, expanel, exmix,exstatlink,exbiolink and extidy.

License GPL (>= 3)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.2

Imports httr,vroom,ggplot2,readxl,grid,gridExtra,writexl,ddpcr,GGally,gt

NeedsCompilation no

Author Bin Wang [aut, cph, cre],
Mingliang Fang [aut],
Changxin Lan [aut],
Menyuan Ren [aut],
Guohuan Zhang [aut],
Ning Gao [aut],
Yanqiu Feng [aut],
Weinan Lin [aut],
Yuting Wang [aut]

Maintainer Bin Wang <binwang@pku.edu.cn>

R topics documented:

BioLink *Build the biological link*

Description

Build the biological link between the exposures and diseases

Usage

```
BioLink(PID, OutPath ="default", Mode, ChemCas="default", ChemInchikey= "default",
  DiseaseID= "default", MetabolomeID= "default", MetBiospec= "blood", ProteomeID= "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitBioLink.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Mode	chr. Method to build the biological link between exposures and diseases. Available options include "PPI" (i.e., protein-protein interaction) and "GO" (i.e., gene ontoloty).
ChemCas	chr. CAS Registry Number of chemicals. Default means using the values in the input data file. Users can also copy the part of them by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "7440-43-9,333-41-5,20461-54-5".
ChemInchikey	chr. InChiKey serial number of chemicals. Default means using the values in the input data file. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "WABPQHHGFIMREM-UHFFFAOYSA-N,BTAGRXWGMYTPBY-UHFFFAOYSA-N,IAKOZHOLGAGEJT-UHFFFAOYSA-N".
DiseaseID	chr. ID of the concerned diseases. Both IDs from OMIM (e.g., OMIM:220100) and MESH (e.g., MESH:C536409) are accepted. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "OMIM:244200,MESH:C536409,OMIM:181500".
MetabolomeID	chr. KEGG entry number of metabolites. Default means using the values in the input data file. Users can also copy the part of them by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "C00022,C00117,C00794".
MetBiospec	chr. Biological sample matrix for the metabolome analysis. Options include "Blood" and "Urine".
ProteomeID	chr. Protein ID. Both IDs of Ensembl and UniProt are accepted. Default means using the values in the input data file. Users can also copy the part of them by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "Q9Y3X0,Q8N5I3, ENSP00000000233,ENSP00000000412".

Value

A list object containing the edges and nodes of the biological link.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res = InitBioLink()
   res1 = LoadBioLink(PID = res$PID, UseExample = "example#1")
   res2 = ConvToExpoID(PID = res$PID)
   res3 = BioLink(PID = res$PID, OutPath="default", Mode = "PPI", ChemCas = "default",
   ChemInchikey = "default",DiseaseID = "default",MetabolomeID = "default",
   MetBiospec = "blood", ProteomeID = "default")
```

ConvToExpoID *Convert different IDs to the unified ExposomeX IDs*

Description

Convert the IDs of exposure, chemicals, metabolites, or proteins to the unified ExposomeX ID, i.e., unified identifier in ExposomeX platform

Usage

```
ConvToExpoID(PID, OutPath="default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitBioLink.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set

Value

A data frame containing the converted ID information

Author(s)

Mingliang Fang, Weinan Lin, Bin Wang (corresponding author)

Examples

```
res = InitBioLink()
   res1 = LoadBioLink(PID = res$PID, UseExample = "example#1")
   res2 = ConvToExpoID(PID = res$PID)
```

`CrosAsso` *Association analysis*

Description

Association analysis for cross-sectional data.

Usage

```
CrosAsso(PID, OutPath = "default", EpiDesign = "cross.sectional",
   VarsY, VarsX = "all.cx",VarsN = "single.factor", VarsSel = FALSE, VarsSelThr = 0.1,
    IncCova = TRUE, Family ,RepMsr = FALSE, Corstr = "ar1")
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by ExpoCros
`OutPath`	chr. Output file directory, e.g., "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
`EpiDesign`	chr. Epidemiological design of the study, including "cohort" "case.control" and "cross.sectional". It doesn't affect the modeling, but the format of the output file. For the three designs, the effect values are usually indicated by RR (relative risk) of cohort, OR (odds ratio) of case-control, and beta value of cross-sectional.
`VarsY`	chr. Outcome variable used for modeling. Only one variable can be entered.
`VarsX`	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
`VarsN`	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
`VarsSel`	lgl. T (or TRUE) and F (or FALSE). Whether to select the significant variable for the final model. Available options.
`VarsSelThr`	num. If "VarsSel" = TRUE, provide the selection threshold of the P-value. three values can be chosen, i.e. 0.05, 0.1, and 0.2.
`IncCova`	lgl. T (or TRUE) and F (or FALSE). Whether to include the covariate selected in the function "FindCovaCros"
`Family`	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable
`RepMsr`	lgl. T (or TRUE) and F (or FALSE). Whether existing repeated measurement of the subjects. Available options.
`Corstr`	chr. If "RepMsr" = TRUE, the generalized estimating equations (GEE) will be used. For GEE, three correlation structure options are "exchangeable" "ar1" "unstructured".

Value

A list containing the association analysis results.

Author(s)

Bin Wang

Examples

```
res <- InitCros()
   res1 = LoadCros(PID = res$PID, UseExample = "example#1")
   res2 = CrosAsso(PID=res$PID, EpiDesign = "cohort",
   VarsY = "Y1", VarsX = "X5,X6,X7,X8,X9,X10,X11", VarsN = "single.factor",
   VarsSel = FALSE, VarsSelThr = 0.1, IncCova = TRUE, Family = "gaussian",
   RepMsr = FALSE,Corstr = "ar1")
   FuncExit(PID = res$PID)
```

CrosPred *Build prediction models*

Description

Build prediction models

Usage

```
CrosPred(PID, OutPath = "default", VarsY, VarsX = "all.x",
    PredType = "response", VarsSel = FALSE, VarsSelThr = 0.1, IncCova = TRUE,
    RsmpMethod = "cv", Folds = 5, Ratio = 0.667, Repeats = 5)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoCros
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
PredType	chr. Prediction type of the outcome variable, including "response" for the actual values and "prob" for outcome with binary variable.
VarsSel	lgl. Whether to select the significant variable for the final model. Available options include T (or TRUE) and F (or FALSE).
VarsSelThr	num. If "VarsSel" = TRUE, provide the selection threshold of the P-value. three values can be chosen, i.e. 0.05, 0.1, and 0.2.

IncCova	lgl. Whether to include the covariate selected in the function of "FindCovaCros". Available options include T (or TRUE) and F (or FALSE).
RsmpMethod	chr. Four resampling methods options for internal validation, including "cv" (i.e.,Cross validation) , "loo" (i.e., eave-one-out), "bootstrap", and "holdout".
Folds	num. Folds of Cross-validation resampling. It is ranging 2-10.
Ratio	num. Ratio of Bootstrap resampling. It is ranging 0.4-0.9.
Repeats	num. Number of Bootstrap resampling. It is ranging 2-20.

Value

A list containing the prediction performance evaluation.

Author(s)

Bin Wang

Examples

```
res <- InitCros()
   res1 = LoadCros(PID = res$PID, UseExample = "example#1")
   res2 = CrosPred(PID=res$PID, VarsY = "Y1", VarsX = "X5,X6,X7,X8,X9,X10,X11",
   PredType = "response", VarsSel = FALSE, VarsSelThr = 0.1, IncCova = FALSE,
   RsmpMethod = "cv", Folds = 5, Ratio = 0.667, Repeats = 5)
   FuncExit(PID = res$PID)
```

DelMiss *Delete variables with missing values*

Description

Whether to delete missing variables with low variance. The default option is "yes". If skipped, it may result in failure during modeling.

Usage

```
DelMiss(PID, OutPath = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.

Value

An R6 class object containing the variable(s) without missing values.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
    res1 = LoadTidy(PID=res$PID, UseExample="example#1")
    res2 = DelMiss(PID=res$PID)
    FuncExit(PID = res$PID)
```

DelNearZeroVar *Delete variables with low variance*

Description

Whether to delete variables with low variance. The default option is "yes". If skipped, it may result in failure to build models.

Usage

```
DelNearZeroVar(PID, OutPath = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.

Value

An R6 class object containing the variable(s) with acceptable variance.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
    res1 = LoadTidy(PID=res$PID, UseExample="example#1")
    res2 = DelNearZeroVar(PID=res$PID)
    FuncExit(PID = res$PID)
```

ExpoAbbr *Explain the abbreviations*

Description

Explain the abbreviations in ExposomeX platform

Usage

```
ExpoAbbr(PID, OutPath = "default", Keys)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoDB
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Keys	chr. Any keywords to search. "default" means using the values . in the input data file. Users can also enter the keywords here. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "EpiDesign,Cros,Cohort".

Value

A data frame

Author(s)

Bin Wang

Examples

```
res <- InitDb()
   res1 = LoadDb(PID = res$PID, UseExample = "example#1")
   res2 = ExpoAbbr(PID=res$PID, Keys = "default")
   FuncExit(PID = res$PID)
```

ExpoAnno *Annotate the non-targeted features*

Description

Annotate the non-targeted features from high-resolution mass spectrometry

Usage

```
ExpoAnno(PID, OutPath = "default", MassToCharge, AdductPos,
   AdductNeg, Accuracy = 1)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoDB
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
MassToCharge	chr. Mass to charge ratio (m/z). It ranges 50-1000. If "default", the values in the input data file are used. Users can also enter the keywords here. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "150,200,210".
AdductPos	chr. Adducts collected in the positive mode. If "default", the values in the input data file are chosen. Users can also enter the keywords here. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "M+H+Na,M+2ACN+2H,M+DMSO+H". All the positive abducts are M+3ACN+2H,M+ACN+H,M+NH4,M+2ACN+2H,M+2H,M+3H,M+3Na, "M+2Na-H,M+ACN+Na,M+H+Na,M+2K-H,M+H+NH4,2M+K,M+K,2M+NH4,2M+Na, M+2Na,M+DMSO+H,M+2ACN+H,M+IsoProp+Na+H,M+2H+Na,M+ACN+2H,M+H, 2M+H,M+CH3OH+H,M+H+2Na,M+Na",2M+ACN+H,2M+ACN+Na,M+IsoProp+H,M+H+K"
AdductNeg	chr. Adducts collected in the negative mode. If "default", the values in the input data file are chosen. Users can also enter the keywords here. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "M+Hac-H,M-2H,M-H2O-H". All the negative abducts are "M+FA-H,M+Hac-H,M+Br,3M-H,2M+Hac-H,M+K-2H,2M+FA-H, M-H,M-H2O-H,M+Na-2H,M-2H,M+TFA-H,M+Cl,M-3H,2M-H".
Accuracy	num. Upper limit of accuracy to match the target molecular.

Value

A data frame

Author(s)

Bin Wang

Examples

```
res <- InitDb()
   res1 = LoadDb(PID = res$PID, UseExample = "example#1")
   res6 = ExpoAnno(PID=res$PID, MassToCharge = "default", AdductPos = "all",
   AdductNeg = "all", Accuracy = 5)
   FuncExit(PID = res$PID)
```

ExpoConv *Convert keywords*

Description

Convert the keywords from different sources in ExposomeX platform.

Usage

```
ExpoConv(PID, OutPath = "default", From, To, Keys)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoDB
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
From	chr. Choose the search range of the convert direction for the concerned keywords (from -> to). Options include "chemical" <-> "cas.rn" (Chemical name <-> CAS Registry Number), "inchikey" <-> "chemical" (InChIKey <-> Chemical name), "metabolite" <-> "kegg.entry" (Metabolite name <-> InChIKey), "kegg.entry" <-> "metabolite" (KEGG Entry ID <-> Metabolite name), "protein" <-> "uniprot" (protein name <-> UniProt ID), "uniprot" <-> "ensembl" (UniProt ID <-> Chemical name), "enzyme" <-> "uniprot" (Enzyme name <-> UniProt ID), "disease" <-> "disease.id" (Disease <-> Disease.id)
To	chr. see "from".
Keys	chr. Any keywords belong to the classes of "from" and "to" to search. "default" means using the values in the input data file. Users can also enter the keywords here. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "7440-43-9,OMIM:619217,zinc,GO:0010942".

Value

A data frame

Author(s)

Bin Wang

Examples

```
res <- InitDb()
   res1 = LoadDb(PID = res$PID, UseExample = "example#1")
   res2 = ExpoConv(PID=res$PID, From = "chemical", To = "cas.rn", Keys = "default")
   FuncExit(PID = res$PID)
```

`ExpoDict` *Explain keywords*

Description

Explain the keyword in ExposomeX platform

Usage

```
ExpoDict(PID, OutPath = "default", Class, Keys)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by ExpoDB
`OutPath`	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
`Class`	chr. Choose the search range of the concerned keywords. Options include "chemical" (Some exposure factors may be a mixture of chemicals, e.g. PM2.5, tobacco smoking. As chemicals account for the majority, we use "Chemical" for convenience), "metabolite" (the chemicals used for metabolome analysis in the KEGG database), "protein", "enzyme" (referring in particular to the enzymes in the KEGG database), "disease", "GO" (gene ontology), and "ion.adduct" (the ion adducts in the liquid chromatograph-mass spectrometry).
`Keys`	chr. Any keywords to search. "default" means using the values in the input data file. Users can also enter the keywords here. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "7440-43-9,OMIM:619217,zinc,GO:0010942".

Value

A data frame

Author(s)

Bin Wang

Examples

```
res <- InitDb()
   res1 = LoadDb(PID = res$PID, UseExample = "example#1")
   res2 = ExpoDict(PID=res$PID, Class = "GO", Keys = "default")
   FuncExit(PID = res$PID)
```

ExpoNexus *Find the nexuses between keywords*

Description

Find the nexuses between the keywords in ExposomeX platform. Nexus direction from keywords A (class A) to keywords B (class B)

Usage

```
ExpoNexus(PID, OutPath = "default", ClassA, ClassB,
   KeysA = "default", KeysB = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoDB
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
ClassA	chr. Find the nexuses between the keywords in ClassA and ClassB. Options include "chemical" <-> "protein", "chemical" <-> "GO", "protein" <-> "protein", "disease" <-> "GO", "protein" <-> "disease"
ClassB	chr. See "ClassA".
KeysA	chr. The lowercases of name, alias, and ID of chemical, metabolite, protein, and enzyme are all accepted. e.g., "7440-43-9,OMIM:619217,zinc,GO:0010942".
KeysB	chr. See "KeysA".

Value

A data frame

Author(s)

Bin Wang

Examples

```
res <- InitDb()
   res1 = LoadDb(PID = res$PID, UseExample = "example#1")
   res2 = ExpoNexus(PID=res$PID, ClassA = "chemical", ClassB = "protein",
   KeysA = "default", KeysB = "default")
   FuncExit(PID = res$PID)
```

FindCovaCros *Find covariates*

Description

Find covariates

Usage

```
FindCovaCros(PID, OutPath = "default", VarsY, VarsC_Prior = "default",
    VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoCros
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsC_Prior	chr. Potential covariates needing further statistical test. The default value is all covariate variables listed in the data file.
VarsC_Fixed	chr. Covariate variables fixed in the model by users.
Method	chr. Methods for screening the covariates, including two options, i.e. "single.factor" and "two.stage".
Thr	num. Threshold of the P-value for screening the covariates. It is ranging 0.05-0.25. The default value is 0.1.

Value

A list containing the selected covariates.

Author(s)

Bin Wang

Examples

```
res <- InitCros()
   res1 = LoadCros(PID = res$PID, UseExample = "example#1")
   res2 = FindCovaCros(PID=res$PID, VarsY = "Y1",
   VarsC_Prior = "default", VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
   FuncExit(PID = res$PID)
```

FindCovaMix *Find covariates*

Description

Find covariates

Usage

```
FindCovaMix(PID, OutPath = "default", VarsY, VarsC_Prior = "default",
    VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoMixEffect
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsC_Prior	chr. Potential covariates needing further statistical test. The default value is all covariate variables listed in the data file.
VarsC_Fixed	chr. Covariate variables fixed in the model by users.
Method	chr. Methods for screening the covariates, including two options, i.e. "single.factor" and "two.stage".
Thr	num. Threshold of the P-value for screening the covariates. It is ranging 0.05-0.25. The defaults value is 0.1.

Value

A list containing the selected covariates.

Author(s)

Bin Wang

Examples

```
res <- InitMix()
   res = LoadMix(PID = res$PID, UseExample = "example#1")
```

FindCovaNta *Find covariates*

Description

Find covariates

Usage

```
FindCovaNta(PID, OutPath = "default", VarsY, VarsC_Prior = "default",
    VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsC_Prior	chr. Potential covariates needing further statistical test. The default value is all covariate variables listed in the data file.
VarsC_Fixed	chr. Covariate variables fixed in the model by users.
Method	chr. Methods for screening the covariates, including two options, i.e. "single.factor" and "two.stage".
Thr	num. Threshold of the P-value for screening the covariates. It is ranging 0.05-0.25. The defaults value is 0.1.

Value

A list containing the selected covariates.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = FindCovaNta(PID=res$PID, VarsY = "Y1", VarsC_Prior = "default",
   VarsC_Fixed = "C2", Method = "single.factor", Thr = 0.1)
```

`FindCovaPanel` *Find covariates*

Description

Find covariates

Usage

```
FindCovaPanel(PID, OutPath = "default", VarsY, VarsC_Prior = "default",
    VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by ExpoPanel
`OutPath`	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
`VarsY`	chr. Outcome variable used for modeling. Only one variable can be entered.
`VarsC_Prior`	chr. Potential covariates needing further statistical test. The default value is all covariate variables listed in the data file.
`VarsC_Fixed`	chr. Covariate variables fixed in the model by users.
`Method`	chr. Methods for screening the covariates, including two options, i.e. "single.factor" and "two.stage".
`Thr`	num. Threshold of the P-value for screening the covariates. It is ranging 0.05-0.25. The defaults value is 0.1.

Value

A list containing the selected covariates.

Author(s)

Bin Wang

Examples

```
res <- InitPanel()
   res1 = LoadPanel(PID = res$PID, UseExample = "example#1")
   res2 = FindCovaPanel(PID=res$PID, VarsY = "Y1",
   VarsC_Prior = "default", VarsC_Fixed ="C1", Method = "single.factor", Thr = 0.1)
```

FindCovaSurv *Find covariates*

Description

Find covariates

Usage

```
FindCovaSurv(PID, OutPath = "default", TimeY, EventY, VarsC_Prior = "default",
    VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoCros
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TimeY	chr. Outcome variable of survival time used for modelling. Only one variable can be entered.
EventY	chr. Outcome variable of status used for modelling. Only one variable can be entered.
VarsC_Prior	chr. Potential covariates needing further statistical test. The default value is all covariate variables listed in the data file.
VarsC_Fixed	chr. Covariate variables fixed in the model by users.
Method	chr. Methods for screening the covariates, including two options, i.e. "single.factor" and "two.stage".
Thr	num. Threshold of the P-value for screening the covariates. It is ranging 0.05-0.25. The default value is 0.1.

Value

A list containing the selected covariates.

Author(s)

Changxin Lan, Bin Wang(corresponding author)

Examples

```
res <- InitSurv()
   res1 = LoadSurv(PID = res$PID, UseExample = "example#1")
   res3 = FindCovaSurv(PID=res$PID, TimeY = "Y1", EventY= 'Y2',
   VarsC_Prior = "default", VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
   FuncExit(PID = res$PID)
```

FuncExit *End the module analysis*

Description

End the module analysis

Usage

```
FuncExit(PID)
```

Arguments

PID	chr. Program ID. It should be the same with the PID generated by initial functions.

Value

Exit status

Author(s)

Bin Wang (corresponding author)

Examples

```
res = InitTidy()
   res1 = LoadTidy(PID=res$PID, UseExample="example#1")
   res2 = DelNearZeroVar(PID=res$PID)
   FuncExit(PID = res$PID)
```

ImptM *Estimate the importance of mediators*

Description

Estimate the importance of mediators.

Usage

```
ImptM(PID, OutPath, VarsY, VarsX = "default", VarsC = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
VarsY	chr. The outcome variable. Either continuous, binary or count type is permitted.
VarsX	chr. Exposure variables included in estimation procedure. When "default" is specified, all exposure variables in the data will be used. The shirnkaged exposure variables are also permitted. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "X1,X2,X3".
VarsC	chr. Covariates included in estimation procedure. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".

Details

ImptM function estimates the importance of mediators among given exposures. The estimation procedure is mainly realized by the hima function in HIMA package. ImptM also provides another evaluation method by the bama function in bama package. The users can search for these packages for further information.

Value

A list containing four elements where the raw estimation results as well as the tidy tables for display are stored. The elements of that list include:

1. "MedtImptM_all": the raw estimation result where the importance of each mediator was evaluated among whole mediators.
2. "MedtImptM_list": the raw estimation result where the importance of each mediator was evaluated among each mediator group.
3. "MedtImptM_table1": the tidy table for MedtImptM_all and MedtImptM_list, where significant estimations (q value <0.2) are expressed with an asterisk mark (*).
4. "MedtImptM_table2": the tidy table for MedtImptM_all and MedtImptM_list, where only significant estimations (q value <0.2) are displayed in this table. MedtImptM_table2 is as same content as MedtImptM_table1.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
res1 <- LoadMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
VocaPath = NULL)
res2 <- XMlists(PID = res$PID)
res3 <- ImptM(PID = res$PID, VarsY = "Y1",
VarsX = "default", VarsC = "default")
res4 <- ImptM(PID = res$PID, VarsY = "Y1",
```

```
VarsX = "X1,X2,X3", VarsC = "C1")
```

`InitBioLink` *Initialize ExpoBioLink module*

Description

Initialize ExpoBioLink module analysis. It can generate an R6 class object integrating all the analysis information

Usage

```
InitBioLink()
```

Details

ExpoBioLink module is designed to find the biological relationships between exposure factors and health outcome. This module adopts the most frequently-used and authoritative databases, e.g., T3DB, CTD, ToxCast, StringDB, STITCH, KEGG, and GO.

Value

An R6 class object.

Author(s)

Mingliang Fang, Bin Wang,(corresponding author)

Examples

```
res <- InitBioLink()
```

`InitCros` *Initialize ExpoCros module*

Description

Initialize ExpoCros module analysis. It can generate an R6 class object.

Usage

```
InitCros()
```

Details

ExpoCros module was designed to analyze the cross-sectional data from exposome-wide association study (EWAS). This data structure can be obtained from the epidemiological designs of cross-section, case-control, and cohort.

Value

An R6 class object.

Author(s)

Bin Wang

Examples

```
res <- InitCros()
    FuncExit(PID = res$PID)
```

`InitDb` *Initialize ExpoDB module*

Description

Initialize ExpoDB module analysis. It can generate an R6 class object.

Usage

```
InitDb()
```

Details

ExpoDB module is designed as a convenient tool to explore the data, as well as facilitating to find the biological relationship between exposure and diseases from the perspective of bioinformatics. This module adopts the most frequently-used and authoritative databases, e.g., T3DB, CTD, ToxCast, StringDB, STITCH, KEGG, and GO.

Value

An R6 class object.

Author(s)

Bin Wang (corresponding author)

Examples

```
res <- InitDb()
   FuncExit(PID = res$PID)
```

InitMedt *Initialize ExpoMediation module*

Description

Initialize ExpoMediation module analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitMedt()
```

Details

InitMedt uses R6 package to generate an R6 class object where parameters to be used for the following mediation module program are initialized and save in that object. All executed function codes in the ExpoMediaiton packaged will be recorded in the form of log text in that object. Furthermore, a program ID (i.e., PID) is randomly created for the users to identify their own program.

Value

An R6 class object.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
```

InitMeta *Initialize ExpoMeta Module*

Description

Initialize ExpoMeta module, the first step to start ExpoMeta Module.

Usage

```
InitMeta()
```

Details

ExpoMeta module mainly provides users preliminary information retrieval and screening for meta-analysis. Run InitMeta to get a unique PID for following steps.

Value

An R6 class object.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitMeta()
```

`InitMix` *Initialize ExpoMixEffect module*

Description

Initialize ExpoMixEffect module analysis. It can generate an R6 class object.

Usage

```
InitMix()
```

Details

ExpoMixEffect module is designed to analyze mixture effect of the various exposure factors. It mainly aims to screen the representative features with high contribution to the health outcome, as well as their potential interaction effect.

Value

An R6 class object.

Author(s)

Bin Wang (corresponding author)

Examples

```
res <- InitMix()
```

InitMO *Initialize ExpoMultiomics module*

Description

Initialize ExpoMultiomics analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitMO()
```

Details

ExpoMultiomics module is designed to integrate the multi-omic data to predict the incidence risk. It mainly aims to construct various stacked generalization(SG) models to predict the probability of outcome incidence, as well as providing the statistical explanation. In addition, the module can provide visualization plots with high quality of the final calculation results to make it easier for users to understand.

Value

An R6 class object.

Author(s)

Guohuan Zhang, Yuting Wang, Bin Wang (corresponding author)

Examples

```
res <- InitMO()
```

InitNTA *Initialize ExpoNTA module*

Description

Initialize ExpoNTA module analysis. It can generate an R6 class object.

Usage

```
InitNTA()
```

Details

ExpoNontarget module is designed to conduct the analysis of the features from the high-resolution mass spectrometry. It mainly aims to screen and annotate the significant features associated with the health outcomes.

Value

An R6 class object.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
```

`InitPanel` *Initialize ExpoPanel module*

Description

Initialize ExpoPanel module analysis. It can generate an R6 class object.

Usage

```
InitPanel()
```

Details

ExpoPanel module is designed to conduct the analysis of the panel data. It mainly aims to evaluate the associations between exposure factors and the health outcome.

Value

An R6 class object.

Author(s)

Bin Wang (corresponding author)

Examples

```
res <- InitPanel()
```

InitStat *Initialize ExpoStat Module*

Description

The first step to start ExpoStat Module

Usage

```
InitStat()
```

Value

An R6 class object.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStat()
```

InitStatLink *Initialize ExpoStatLink module*

Description

Initialize ExpoStatLink module analysis. It can generate an R6 class object.

Usage

```
InitStatLink()
```

Details

ExpoStatLink module is designed to find the statistical relationships between exposure factors and health outcome.

Value

An R6 class object.

Author(s)

Bin Wang

Examples

```
res <- InitStatLink()
```

`InitSurv` *Initialize ExpoSurvival module*

Description

Initialize ExpoSurvival module analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitSurv()
```

Details

InitSurv uses R6 package to generate an R6 class object where parameters to be used for the following mediation module program are initialized and save in that object. All executed function codes in the ExpoSurvival packaged will be recorded in the form of log text in that object. Furthermore, a program ID (i.e., PID) is randomly created for the users to identify their own program.

Value

An R6 class object.

Author(s)

Changxin Lan, Bin Wang(corresponding author)

Examples

```
res <- InitSurv()
```

`InitTidy` *Initialize ExpoTidy module*

Description

Initialize ExpoTidy module analysis. It can generate an R6 class object.

Usage

```
InitTidy()
```

Details

It is designed to tidy the data for the target model analysis.

Value

An R6 class object.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
   FuncExit(PID = res$PID)
```

`InitViz` *Initialize ExpoViz module*

Description

Initialize ExpoViz module analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitViz()
```

Details

ExpoViz module is designed for the data visualization of different statistical and biological analyses in a user friendly and easy way, including four typical classes of visualization.

Value

An R6 class object.

Author(s)

Ning Gao, Bin Wang (corresponding author)

Examples

```
res <- InitViz()
```

LoadBioLink *Load data file for BioLink module*

Description

Load data file for BioLink module

Usage

```
LoadBioLink(PID, UseExample = "default", DataPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitBioLink.
UseExample	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input file directory, e.g. "D:/test/eg_biolink.xlsx". It should be noted that the slash symbol is "/", not "\"

Value

An R6 class object containing the input data.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitBioLink()
   res = LoadBioLink(PID = res$PID, UseExample = "example#1")
```

LoadCros *Load data file for ExpoCros module*

Description

Load data file for ExpoCros module.

Usage

```
LoadCros(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoCros
UseExample	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_expocros.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_expocros.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitCros()
   res1 = LoadCros(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

LoadDb *Load data file for ExpoDB module*

Description

Load data file for ExpoDB module

Usage

```
LoadDb(PID, UseExample = "default", DataPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoDB
UseExample	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitDb()
   res1 = LoadDb(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

LoadMedt *Load data file for Mediation module*

Description

Load data file for Mediation module

Usage

```
LoadMedt(PID, UseExample = "default", DataPath = NULL, VocaPath = NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMedt.
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input data file directory, e.g. "D:/test/eg_Medt_data.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input vocabulary file directory, e.g. "D:/test/eg_Medt_data.xlsx". It should be noted that the slash symbol is "/", not "\".

Details

LoadMedt function loads the data file and the vocabulary file into the R6 object that InitMedt created. Noted that there are several data format requirments for the data and vocabulary file. For data file, the first three column must be named as "SampleID", "SubjectID" and "Group" in sequence. The "Gourp" variable should be a character variable to category data into two groups: "train" and "test" group. Outcome variables should be named as "Y*". e.g., Y1, Y2, Y3... Similarly, exposure variables should be named as "X*". e.g., X1, X2, X3..., and mediator variables should be named as "M*". e.g., M1, M2, M3... For vocabulary file, the first column should be a character variable named "SerialNo" indicating the names of outcome, exposure and mediator variables in the data file. These names should be consistent with the variable names in data file. The second column should be a character variable named "FullName" indicating the full names (labels) of the variables. The third column should be a character variable named "SubgroupName" indicating the groups the exposure or mediator variables belong to.

Value

An R6 class object containing the input data and vocabulary file.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
   res1 <- LoadMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
```

LoadMeta *Load Data for ExpoMeta Module*

Description

Upload local data file for ExpoMeta Module.

Usage

```
LoadMeta(PID, UseExample = "default", DataPath = NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMeta.
UseExample	chr. A character indicates whether uses example data for analyses, available option include "example#1" for using example data1 and "default" for using data uploaded.
DataPath	chr. Input file directory, e.g. "D:/test/eg_meta.xlsx". It should be noted that the slash symbol is "/", not "\".

Details

After initializing ExpoMeta module, the second step is to upload local data file for ExpoMeta Module. Set param "UseExample = 'example#1'" to use example data1. LoadMeta can only run successfully after successfully running InitMeta. Please attention, PID must be got from the return result of InitMeta().

Value

An R6 class object containing the input data.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitMeta()
   res1 <- LoadMeta(PID = res$PID, UseExample = "example#1", DataPath = NULL)
```

LoadMix *Load data file for ExpoMixEffect module*

Description

Load data file for ExpoMixEffect module

Usage

```
LoadMix(PID, UseExample = "default", DataPath = NULL, VocaPath = NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoMixEffect
UseExample	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitMix()
   res = LoadMix(PID = res$PID, UseExample = "example#1")
```

LoadMO *Load data file for multiomics module*

Description

Upload data file for multiomics module.

Usage

```
LoadMO(PID, UseExample= "default", DataPath, VocaPath)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMo.
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input file directory, e.g. "D:/test/eg_data.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input file vocabulary, e.g. "D:/test/eg_voca.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Guohuan Zhang, Yuting Wang, Bin Wang (corresponding author)

Examples

```
 res <- InitMO()
res <- LoadMO(PID=res$PID, UseExample = "example#1", DataPath = NULL, VocaPath = NULL)
```

LoadNTA *Load data file for ExpoNTA module*

Description

Load data file for ExpoNTA module

Usage

```
LoadNTA(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
UseExample	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res = LoadNTA(PID = res$PID, UseExample = "example#1")
```

LoadPanel *Load data file for ExpoPanel module*

Description

Load data file for ExpoPanel module

Usage

```
LoadPanel(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoPanel
UseExample	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitPanel()
   res = LoadPanel(PID = res$PID, UseExample = "example#1")
```

LoadStat *Upload data file for ExpoStat Module*

Description

Upload data file for ExpoStat Module

Usage

```
LoadStat(PID, UseExample = "default", DataPath = NULL, VocaPath = NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStat.
UseExample	chr. Whether uses example data for analyses, available option include "example#1" for using example data1 and "default" for using data.
DataPath	chr. Input file directory, e.g. "D:/test/expostat_data.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input file directory, e.g. "D:/test/expostat_voca.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

A list object containing imported data.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStat()
   res1 = LoadStat(PID = res$PID, UseExample = "example#1")
```

LoadStatLink *Load data file for ExpoStatLink for module*

Description

Load data file for ExpoStatLink module

Usage

```
LoadStatLink(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoStatLink
UseExample	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitStatLink()
   res = LoadStatLink(PID = res$PID, UseExample = "example#1")
```

LoadSurv *Load data file for Survival module*

Description

Load data file for Survival module

Usage

```
LoadSurv( PID, UseExample = "default", DataPath=NULL,VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitSurv.
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input data file directory, e.g. "D:/test/eg_Surv_data.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input vocabulary file directory, e.g. "D:/test/eg_Surv_voca.xlsx". It should be noted that the slash symbol is "/", not "\".

Details

LoadSurv function loads the data file and the vocabulary file into the R6 object that InitSurv created. Noted that there are several data format requirments for the data and vocabulary file. For data file,the first three columns should be named as "SampleID", "SubjectID", and "Group", respectively. For the "Group" variable, only two values can be used, i.e. "train" and "test". If there is no data for test, all values should be set as "train". For outcome variables, their initials must be set as "Y" and serialized by adding Arabic numerals if needed, e.g., Y1, Y2, Y3. In this module, the survival time (Y1) and status (Y2) must be provided.For exposure variables, their initials must be set as "X" and serialized by adding Arabic numerals if needed, e.g., X1, X2, X3. For covariate variables, their initials must be set as "C" and serialized by adding Arabic numerals if needed, e.g., C1, C2, C3. It should be noted the covariates are not required if users don't have. For vocabulary file, the first two columns must be named as "SerialNo" and "FullName", respectively. The list of SerialNo of outcomes, exposure, and covariates should be the same with the column names of "Data file". The list of the FullName is prepared as users' like.

Value

An R6 class object containing the input data and vocabulary file.

Author(s)

Changxin Lan, Bin Wang(corresponding author)

Examples

```
res <- InitSurv()
   res1 <- LoadSurv(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
```

LoadTidy *Load data file for ExpoTidy module*

Description

Load data file for ExpoCros module

Usage

```
LoadTidy(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoCros
UseExample	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitTidy()
   res = LoadTidy(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

LoadViz *Load data file for ExpoViz module*

Description

Load data for visualization.

Usage

```
LoadViz(PID, UseExample = "example#1",DataPath=NULL,VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitViz.
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input data file directory, e.g. "D:/test/eg_expoviz_data.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input vocabulary file directory, e.g. "D:/test/eg_expoviz_voca.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Ning Gao, Bin Wang (corresponding author)

Examples

```
res = InitViz()
    res1 = LoadViz(PID = res$PID, UseExample = "example#1")
```

`MetaEffect` *Pool Effect Value*

Description

Pool effect value in our meta database besed on chemical ID and disease ID

Usage

```
MetaEffect(PID, OutPath = "default", CID = "default", DID = "default")
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitMeta.
`OutPath`	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set. It should be noted that the slash symbol is "/", not "\".
`CID`	chr. "default" or a chemical ID character (separate different values by ","). If "default", the function will use the Chemical_ID values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the chemical ID in the character instead. Chemical_ID refers to the target chemical ID which can be inchikey (eg. JIAARYAFYJHUJI-UHFFFAOYSA-L), cas.rn (eg. 7784-42-1) or our EXC ID (eg. EX:C01631).
`DID`	chr. "default" or a disease ID character (separate different values by ","). If "default", the function will use the Disease_ID values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the disease ID in the character instead. Disease_ID refers to the target disease ID which can be MESH ID (format like MESH:D006973), OMIM ID (format like OMIM:182940) or our EXD ID (eg. EX:D16243)

Details

MetaEffect provides the functions of effect value pooling. In the publishd papers, what is the effect value between X and Y? This questions can be solved by MetaEffect function. It can provide the combined results of fixed effect model and random effect model. (It can only search the papers available in ExpoMeta database DB_Meta) Please attention, PID must be got from the return result of InitMeta(). MetaEffect can only run successfully after successfully running InitMeta and LoadMeta functions.

Value

A list object containing forest plots.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitMeta()
   res1 <- LoadMeta(PID = res$PID,
   UseExample = "example#1",
   DataPath = NULL)
   res2 <- MetaEffect(PID=res$PID,
   OutPath = "default",
   CID = "default",
   DID = "default")
```

MetaRefer *Search or Download Articles' Main Information*

Description

Search or download articles' main information based on keywords.

Usage

```
MetaRefer(PID, OutPath = "default", Mode = "search",
    VarX = "default", VarY = "default", VarM = "default", YearFrom = "default",
    YearEnd = "default", PMID = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMeta.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set. It should be noted that the slash symbol is "/", not "\".
Mode	chr. Two modes are provided. "Search" for paper retrieval by keywords VarX/VarY/VarM/YearFrom/Yearl and "Download" for downloading information (main information only) for specified PMID.

VarX	chr. "default" or a chemical names character (separate different values by ","). If "default", the function will use the VarX values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the chemical names in the character instead.
VarY	chr. "default" or a disease names character (separate different values by ","). If "default", the function will use the VarY values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the disease names in the character instead.
VarM	chr. "default" or a mediating factor names character (separate different values by ","). If "default", the function will use the VarM values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the mediating factor names in the character instead.
YearFrom	chr. "default" or a year character (separate different values by ","). Limits the time range searched for "search" mode. If "default", the function will use the YearFrom values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the YearFrom in the character instead.
YearEnd	chr. "default" or a year character (separate different values by ","). Limits the time range searched for "search" mode. If "default", the function will use the YearEnd values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the YearEnd in the character instead.
PMID	chr. "default" or a PMID character (separate different values by ","). Refers to the papers PMID for "download" mode. If "default", the function will use the PMID values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the PMIDs in the character instead.

Details

MetaRefer provides the functions of paper retrieval and relevance sorting, returning the information to the user based on keywords. Please attention, PID must be got from the return result of InitMeta(). MetaRefer can only run successfully after successfully running InitMeta and LoadMeta functions.

Value

A list object containing dataframe of articles' information.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitMeta()
   res1 <- LoadMeta(PID = res$PID, UseExample = "example#1",  DataPath = NULL)
   res2 <- MetaRefer(PID = res$PID, OutPath = "default", Mode = "search",
   VarX = "default", VarY = "default", VarM = "default", YearFrom = "default",
   YearEnd = "default", PMID = "default")
```

`MetaReview` *Review Relationship Between Exposure and Outcome*

Description

Review the relationship between exposure and outcome besed on chemical ID and disease ID.

Usage

```
MetaReview(PID, OutPath = "default", CID = "default", DID = "default")
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitMeta.
`OutPath`	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set. It should be noted that the slash symbol is "/", not "\".
`CID`	chr. "default" or a chemical ID character (separate different values by ","). If "default", the function will use the Chemical_ID values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the chemical ID in the character instead. Chemical_ID refers to the target chemical ID which can be inchikey (eg. JIAARYAFYJHUJI-UHFFFAOYSA-L), cas.rn (eg. 7784-42-1) or our EXC ID (eg. EX:C01631).
`DID`	chr. "default" or a disease ID character (separate different values by ","). If "default", the function will use the Disease_ID values in the file loaded by LoadMeta. If a character (separate different values by ","), the function will use the disease ID in the character instead. Disease_ID refers to the target disease ID which can be MESH ID (format like MESH:D006973), OMIM ID (format like OMIM:182940) or our EXD ID (eg. EX:D16243)

Details

MetaReview provides the functions of literature review. In the publishd papers, how many recorded that X is a protective/risky factor for Y? This questions can be solved by MetaReview function. (It can only search the papers available in our database) Please attention, PID must be got from the return result of InitMeta(). MetaReview can only run successfully after successfully running InitMeta and LoadMeta functions.

Value

A list object containing relationship visualization.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitMeta()
   res1 <- LoadMeta(PID = res$PID,
   UseExample = "example#1",
   DataPath = NULL)
   res2 <- MetaReview(PID = res$PID,
   OutPath = "default",
   CID = "default",
   DID = "default")
```

MixBKMR *Build the Bayesian Kernel Machine Regression (BKMR) model*

Description

Build the Bayesian Kernel Machine Regression (BKMR) model

Usage

```
MixBKMR(PID, OutPath = "default", VarsY, VarsX, IncCova,
    Family, Group = F,Iter = 2000, qfixed = 0.5, qsbivar = "default",
    qsoverall = "default", qsdiff = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoMixEffect
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
IncCova	lgl. T (or TRUE) and F (or FALSE). Whether to include the covariate selected in the function "FindCovaMix"
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable
Group	lgl. T (or TRUE) and F (or FALSE). Whether to use group indicators for fitting hierarchical variable selection. If "TRUE", the group name (GroupName) should be provided in the vocabulary data file.
Iter	num. Number of iterations for modeling. The default is 500. For more accurate and stable results, a minimum of 10,000 iteration is recommended.
qfixed	num. Quantile at which to fix the other predictors. The default is 0.5.

qsbivar	chr. Quantiles at which to fix the second variable. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., the default character sequence is "0.1,0.5,0.9"
qsoverall	num. Quantiles at which to calculate the overall risk summary. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., the default character sequence is "0.25,0.30,0.35,0.40,0.45,0.50,0.55,0.60,0.65,0.7
qsdiff	chr. Indicating the two quantiles for computing their effect difference. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., the default character sequence is "0.25,0.75".

Value

A list containing the BKMR analysis results.

Author(s)

Bin Wang

Examples

```
res <- InitMix()
   res1 = LoadMix(PID = res$PID, UseExample = "example#1")
   res2 = MixBKMR(PID = res$PID, VarsY = "Y1", VarsX = "X4,X5,X6,X7,X8,X9,X10",
  IncCova = 'F', Family = "gaussian", Group = 'F',Iter = 2000,qfixed = 0.5,qsbivar = "default",
   qsoverall = "default",qsdiff = "default")
```

MixMLR *Build multiple linear regression (MLR) model*

Description

Build multiple linear regression (MLR) model

Usage

```
MixMLR(PID,OutPath = "default",VarsY,VarsX,IncCova = "F",
    SelMethod = "lasso",PredType = "response",Family)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoMixEffect
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"

IncCova	lgl. T (or TRUE) and F (or FALSE). Whether to include the covariate selected in the function "FindCovaMix"
SelMethod	chr. Method to select the important features to the final model. Options include "stepwise" (stepwise regression), "lasso" (Regularization regression of least absolute shrinkage and selection operator), and "enet" (Regularization regression of elastic net).
PredType	chr. Prediction type of the outcome variable, including "response" for the actual values and "prob" for outcome with binary variable.
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable

Value

A list containing the MLR analysis.

Author(s)

Bin Wang

Examples

```
res <- InitMix()
   res1 = LoadMix(PID = res$PID, UseExample = "example#1")
   res2 = MixMLR(PID = res$PID, VarsY = "Y1", VarsX = "all.x", IncCova = "F",
   SelMethod = "lasso", PredType = "response", Family = "gaussian")
```

MixWQS *Build weighted quantile sum regression (WQS) model*

Description

Build weighted quantile sum regression (WQS) model

Usage

```
MixWQS(PID, OutPath = "default", VarsY, VarsX, IncCova = "F", Family,
    VarStrat = "none",RatioValidat = 0.3,q = 10,b = 100,b1_pos = F,b1_constr = F)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoMixEffect
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.

VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
IncCova	lgl. T (or TRUE) and F (or FALSE). Whether to include the covariate selected in the function "FindCovaMix"
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable.
VarStrat	chr. A factor variable used for stratifying for the model.
RatioValidat	num. Percentage of the dataset to be used to validate the model. If validation = 0 then the test dataset is used as validation dataset too. The default is 0.3.
q	num. Levels for ranking mixture variables, e.g. in quartiles (q = 4), deciles (q = 10), or percentiles (q = 100). The default is 10.
b	num. Number of bootstrap samples used in parameter estimation. The default is 100.
b1_pos	lgl. T (or TRUE) and F (or FALSE). Whether the beta values were positive to derive weights from to build models.
b1_constr	lgl. T (or TRUE) and F (or FALSE). A logial value that determines whether to apply positive (if b1_pos = TRUE) or negative (if b1_pos = FALSE) constraints in the optimization function for the weight estimation.

Value

A list containing the WQS analysis results.

Author(s)

Bin Wang

Examples

```
res <- InitMix()
   res1 = LoadMix(PID = res$PID, UseExample = "example#1")
   res3 = MixWQS(PID=res$PID, VarsY = "Y1", VarsX = "all.x", IncCova = "F",
   Family = "gaussian", VarStrat = "none", RatioValidat = 0.3,
   q = 10, b=100, b1_pos = 'F', b1_constr = 'F')
```

MulOmicsCros *Build multiomics model*

Description

MulOmicsCros function is designed to integrate the multi-omic data to predict the incidence risk. It mainly aims to construct various stacked generalization models to predict the probability of outcome incidence, as well as providing the statistical explanation.

Usage

```
MulOmicsCros(PID, OutPath, OmicGroups,VarsY, VarsC, TuneMethod = "default", TuneNum,
       RsmpMethod, Folds, Ratio, Repeats, VarsImpThr, SG_Lrns)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMo.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
OmicGroups	chr. Groups to be integrated. The groups of outcome and covariates or confounders are not included. Note that separates different learners by "," and without space(e.g. OmicGroups = "immunome,metabolome,proteome").
VarsY	chr. Outcome variable for modelling. Only one variable can be entered.
VarsC	chr. Covariates needing further statistical test. "all.c" option refers to all covariate variables listed in the data file. Users can also select part of them by copying available vars. Note that separates different vars by "," and without space(e.g. VarsC = "C1,C2").
TuneMethod	chr. Method for hyper-parameter autotuning. Options include "default", "random_search", "grid_search", "nloptr"(Non-linear optimization), and "gensa"(Generalized simulated annealing). The "default" option uses the simple training method for parameter optimization of mlr3 package.
TuneNum	num. Upper limit of model tuning times. It should be more than 20 times to search the appropriate parameters, but it takes more time. In theory, more time, better training results.
RsmpMethod	chr. Method for resampling. Options include "cv"(cross validation), "loo"(leave-one-out cross validation), "bootstrap"(bootstrapping), "holdout"(holdout).
Folds	num. Folds for cross validation resampling method. The default value is 5.
Ratio	num. Ratio for "Holdout" resampling method. The default value is 5.
Repeats	num. Repeats for "Bootstrap" resampling method.
VarsImpThr	num. Threshold for feature selection. It refers to the ratio of accumulated importance of all variables of the selected variables for building the final model.
SG_Lrns	chr. Learners for stacked generalization. Options include "lasso", "enet"(Elastic net), "rf"(Random forest), and "xgboost"(Xgboost). One or more arbitrary options can be selected at the same time. Note that separates different learners by "," and without space(e.g. SG_Lrns ="lasso,enet,rf,xgboost").

Details

The calculation time depends on the characteristics of your data, the number of learning methods, and the tuning method. For parameter "TuneMethod", the default option can provide faster calculations but less accurate results than other autotune methods. If you want to train a better model, choose other auto-tune method and increase the number of tuning times.

Value

An R6 class object containing eight elements. The elements of that object include: (1) "Importance": A list containing dataframes that contain the importance of features after modeling a single omic from different omicgroups. (2) "Feature": A list containing dataframes that contain the the coefficients or importance of selected features after modeling a single omic from different learners. (3) "Feature_select": A list containing dataframes that contain the selected features after modeling a single omic from different learners. (4) "ModelStat": A list containing dataframes that contain the r-square value of the single omic model built by different learners. (5) "Prediction_comp": A list containing dataframes that contain the prediction values of the SG model built by different combinations of learners. (6) "SGModel_summary": A dataframe containing the r-square value of the SG model built by different combinations of learners. (7) "NodeNum": The node number generated by different models. (8) "SGplot": A visualized plot for SG model summary.

Author(s)

Guohuan Zhang, Yuting Wang, Bin Wang (corresponding author)

Examples

```
 res <- InitMO()
 res <- LoadMO(PID=res$PID, UseExample = "example#1", DataPath = NULL, VocaPath = NULL)
res2 <- MulOmicsCros(PID=res$PID, OutPath = "default", OmicGroups = "immunome,metabolome,proteome",
VarsY = "Y1", VarsC = "all.c", TuneMethod = "random_search", TuneNum = 5, RsmpMethod = "cv", Folds = 5,
 Ratio = 0.67, Repeats = 5, VarsImpThr = 0.85, SG_Lrns ="lasso,enet,rf,xgboost")
```

NtaAnno *Annotate the non-targeted features*

Description

Annotate the non-targeted features

Usage

```
NtaAnno(PID, OutPath = "default", VarsY, VarsX  = "default",
   VarsN  = "single.factor", FdrCorrect = F, AdductPos = "all", AdductNeg = "all",Accuracy)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"

VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
FdrCorrect	lgl. T (or TRUE) and F (or FALSE). Whether to correct the multiple hypotheses by false positive rate (FDR) method.
AdductPos	chr. Adducts formed in the positive mode using LC-MS. The default is "all", i.e., all the adducts in positive mode will be chosen, including "M+3ACN+2H,M+ACN+H,M+NH4,M+2ACN H,M+ACN+Na,M+H+Na, M+2K-H,M+H+NH4,2M+K,M+K,2M+NH4,2M+Na,M+2Na,M+DMSO+H, M+2H+Na,M+ACN+2H,M+H,2M+H,M+CH3OH+H,M+H+2Na,M+Na,2M+ACN+H,2M+ACN+Na,M
AdductNeg	chr. Adducts formed in the negative mode using LC-MS. The default is "all", i.e., all the adducts in positive mode will be chosen, including "M+FA-H,M+Hac-H,M+Br,3M-H,2M+Hac-H,M+K-2H,2M+FA-H,M-H,M-H2O-H,M+Na-2H, M-2H,M+TFA-H,M+Cl,M-3H,2M-H"
Accuracy	num. Accuracy threshold to match the target compounds. The default is 1 ppm.

Value

A list containing non-target analysis results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = NtaCros(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = TRUE, SelMethod = "all",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, IncCova = "F", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
   res3 = NtaAnno(PID=res$PID,VarsY = "Y1",VarsX = "default",VarsN = "single.factor",
   FdrCorrect = "F",AdductPos = "M+H",AdductNeg = "M-H",Accuracy = 1)
```

NtaCros *Association analysis*

Description

Association analysis for non-targeted data

Usage

```
NtaCros(PID,OutPath = "default", VarsY, VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = T, SelMethod = "all", StepwizeThr = 0.1,
    RF_ImpThr = 0.9, IncCova = F, Family, RepMsr = F, Corstr= "ar1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
FdrCorrect	lgl. T (or TRUE) and F (or FALSE). Whether to correct the multiple hypotheses by false positive rate (FDR) method.
SelMethod	chr. Methods to select the significant features. Options include "stepwise" (multiple linear regress using stepwise algorithm), "lasso" (multiple linear regress using LASSO regularization algorithm), "random.forest" (random forest), and "all" (combination of the above three method)..
StepwizeThr	num. Threshold of the P value for stepwise regression to screen important variables. It ranges 0.05-0.25 with the default value of 0.1.
RF_ImpThr	num. Threshold of the total importance for the variables to a random forest model. It ranges 0.5-1.0 with the default value of 0.9.
IncCova	lgl. T (or TRUE) and F (or FALSE). Whether to include the covariate selected in the step of "FindCovaNta".
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable
RepMsr	lgl. T (or TRUE) and F (or FALSE). Whether existing repeated measurement of the subjects.
Corstr	chr. If "RepMsr" = T, the generalized estimating equations (GEE) will be used. For GEE, three correlation structure options are "exchangeable" "ar1" "unstructured".

Value

A list containing non-target analysis association results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = NtaCros(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
```

```
VarsN = "single.factor", FdrCorrect = T, SelMethod = "all",
StepwizeThr = 0.1, RF_ImpThr = 0.9, IncCova = "F", Family = "gaussian",
RepMsr = "F", Corstr = "ar1")
```

`Pairwise` *Implement pairwise mediation analyses*

Description

Implement pairwise mediation analyses for each pair of exposure and mediator.

Usage

```
Pairwise(PID, OutPath, VarsY, VarsX = "default", VarsM = "default",
    VarsC = "default", Family, Iter = 500)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitMedt.
`OutPath`	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
`VarsY`	chr. The outcome variable. Either continuous, binary or count type is permitted.
`VarsX`	chr. Exposure variables included in pairwise mediation modelling. When "default" is specified, all exposure variables in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "X1,X2,X3".
`VarsM`	chr. Mediator variables included in pairwise mediation modelling. When "default" is specified, all mediator variables in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "M1,M2,M3".
`VarsC`	chr. Covariates included in pairwise mediation modelling. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".
`Family`	chr. The error distribution and link function to be used in the model. Available options include "linear" and "gaussian" for continuous outcome variables, and "binomial" as well as "poisson" options for binary and count outcome variables.
`Iter`	num. The number of iteration times. Default is 500. To get a more stable result, a minimum iteration of 1,000 is recommended. To obtain more digits of P value of parameter estimation, a minimum iteration of 5,000 is suggested.

Details

Pairwise function implements mediation modelling for each pair of exposure and mediator. Given M exposures and N mediators, an exhaustive rule will be executed and M*N pair-wised mediation modelling are fitted. The modelling was realized using mediate function in mediation package.

Value

A list containing one dataframe that contains the pairwise mediation modelling results.

1. "MedtPairWise_Stats": pairwise mediation modelling results.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
res1 <- LoadMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
VocaPath = NULL)
res2 <- XMlists(PID = res$PID)
res3 <- Pairwise(PID=res$PID, VarsY = "Y1",
VarsX = "default", VarsM = "default", VarsC = "default", Family = "linear",
Iter = 500)
res4 <- Pairwise(PID=res$PID, VarsY = "Y1",
VarsX = "X1,X2,X3", VarsM = "M1,M2,M3", VarsC = "C1", Family = "linear",
Iter = 500)
```

PanelAsso *Association analysis of panel data*

Description

Association analysis of panel data

Usage

```
PanelAsso(PID, OutPath = "default", VarsY, VarsX, VarsN = "single.factor",
    VarsRandomIpt = "SubjectID", VarsRandomSlp = "none", IncCova = F)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoPanel
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"

VarsRandomIpt	chr. Random intercept variable for the linear mixed-effect model. The default is "SubjectID".
VarsRandomSlp	chr. Random slope variable for the linear mixed-effect model. The default is "none". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
IncCova	lgl. T (or TRUE) and F (or FALSE). Whether to include the covariate(s) selected in the function "FindCovaPanel"

Value

A list containing the association analysis results.

Author(s)

Bin Wang

Examples

```
res <- InitPanel()
   res1 = LoadPanel(PID = res$PID, UseExample = "example#1")
   res2 = PanelAsso(PID=res$PID, VarsY = "Y1",
    VarsX = "X1,X2,X3,X4,X5,X6,X7,X8,X9,X10,X11,X12", VarsN = "single.factor",
    VarsRandomIpt = "SubjectID", VarsRandomSlp = "none", IncCova = F)
```

RedM *Mediator dimension reduction*

Description

Implement mediator dimension reduction for mediators.

Usage

```
RedM(PID, OutPath, VarsY, VarsX = "default", VarsC = "default",
    Method = "mean", Family, Iter = 500)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
VarsY	chr. The outcome variable. Either continuous, binary or count type is permitted.
VarsX	chr. Exposure variables included in mediation modelling. When "default"(recommended) is specified, all exposure variables in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "X1,X2,X3".

VarsC	chr. Covariates included in mediation modelling. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".
Method	chr. Dimension reduction method. Available options include "mean" and "pdm1"(default). "mean" option is recommended when outcome variable is a binary variable because in some situations error might occur in "pdm1" settings for binary outcome.
Family	chr. The error distribution and link function to be used in the model. Available options include "linear" and "gaussian" for continuous outcome variables, and "binomial" as well as "poisson" options for binary and count outcome variables.
Iter	num. The number of iteration times. Default is 500. To get a more stable result, a minimum iteration of 1,000 is recommended. To obtain more digits of P value of parameter estimation, a minimum iteration of 5,000 is suggested.

Details

RedM function provides two alternative methods for mediator dimension reduction, including sum method as well as pdm1 method via PDM package. By default, all mediators will be included for dimension reduction. Afterwards, mediation models will be built between given exposures as well as shrinkaged mediator variables.

Value

A list containing two elements where the mediator dimension reduction modelling results with exposures are stored. That list include:

1. "MedtRedM_all": a dataframe containing the mediator dimension reduction result where all mediators are considered as one group and that shrinkaged mediator is paired and modelled with each exposure.
2. "MedtRedM_list": a dataframe containing the mediator dimension reduction result where mediators are shrinkaged in their own subgroups, and these shrinkaged mediators are paired and modelled with each exposure.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
res1 <- LoadMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
VocaPath = NULL)
res2 <- XMlists(PID = res$PID)
res3 <- RedM(PID=res$PID, VarsY = "Y1", VarsX = "default",
VarsC = "default", Method = "mean", Family = "linear", Iter = 500)
res4 <- RedM(PID=res$PID, VarsY = "Y1", VarsX = "X1,X2,X3",
VarsC = "C1", Method = "mean", Family = "linear", Iter = 500)
```

RedX *Expoures dimension reduction*

Description

Implement dimension reduction for exposures.

Usage

```
RedX(PID, OutPath, VarsY, VarsC = "default", VarsM = "default",
    Method = "mean", Folds = 10, Family, Iter = 500)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
VarsY	chr. The outcome variable. Either continuous, binary or count type is permitted.
VarsC	chr. Covariates included in mediation modelling. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".
VarsM	chr. Mediator variables included in pairwise mediation modelling. When "default" is specified, all mediator variables in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "M1,M2,M3".
Method	chr. Dimension reduction method. Available options include "gcdnet" and "mean"(default). "mean" option is recommended when outcome variable is a binary variable because in some situations low variance of shrinkaged exposure variable might obtain in "gcdnet" settings for binary outcome.
Folds	num. Number of cross validation for gcdnet method. Default is 10.
Family	chr. The error distribution and link function to be used in the model. Available options include "linear" and "gaussian" for continuous outcome variables, and "binomial" as well as "poisson" options for binary and count outcome variables.
Iter	num. The number of iteration times. Default is 500. To get a more stable result, a minimum iteration of 1,000 is recommended. To obtain more digits of P value of parameter estimation, a minimum iteration of 5,000 is suggested.

Details

RedX function provides two alternative methods for exposure dimension reduction, including sum method as well as adaptive elastic net method via gcdnet package. By default, all exposures will be included for dimension reduction. Afterwards, mediation models will be built between given mediators as well as shrinkaged exposure variables.

Value

A list containing three elements where the exposure dimension reduction variables as well as their mediation modelling results with mediators are stored. That list include:

1. "MedtRedX_ERSall": a dataframe containing the exposure dimension reduction result where all exposures are considered as one group (ERS_All).
2. "MedtRedX_ERSlist": a dataframe containing the exposure dimension reduction results where exposures are shrinkaged in their own subgroups.
3. "MedtRedX_Stats": a dataframe containing the mediation modelling results where the shrinkaged exposure variables were paired with each mediator provided.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
res1 <- LoadMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
VocaPath = NULL)
res2 <- XMlists(PID = res$PID)
res3 <- RedX(PID=res$PID, VarsY = "Y1",
VarsC = "default", VarsM = "default", Method = "mean", Folds = 10,
Family = "linear", Iter = 500)
res4 <- RedX(PID=res$PID, VarsY = "Y1",
VarsC = "C1", VarsM = "M1,M2,M3", Method = "mean", Folds = 10,
Family = "linear",Iter = 500)
```

RedXM *Exposure and mediator dimension reduction*

Description

Implement exposure and mediator dimension reduction. It should be noted that the exposure dimension reduction result has been built by RedX function prior to using it.

Usage

```
RedXM(PID, OutPath, VarsY, VarsC = "default", Method = "mean",
    Family, Iter = 500)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
VarsY	chr. The outcome variable. Either continuous, binary or count type is permitted.
VarsC	chr. Covariates included in mediation modelling. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".
Method	chr. Dimension reduction method. Available options include "mean" and "pdm1"(default). "mean" option is recommended when outcome variable is a binary variable because in some situations error might occur in "pdm1" settings for binary outcome.
Family	chr. The error distribution and link function to be used in the model. Available options include "linear" and "gaussian" for continuous outcome variables, and "binomial" as well as "poisson" options for binary and count outcome variables.
Iter	num. The number of iteration times. Default is 500. To get a more stable result, a minimum iteration of 1,000 is recommended. To obtain more digits of P value of parameter estimation, a minimum iteration of 5,000 is suggested.

Details

RedXM function provides two alternative methods for mediator dimension reduction, including sum method as well as pdm1 method via PDM package. By default, all mediators will be included for dimension reduction. Afterwards, mediation models will be built between shrinkaged exposure variables as well as shrinkaged mediator variables. Prior to using RedXM, make sure that RedX function has been executed to obtain shrinkaged exposure variables.

Value

A list containing two elements where the mediator dimension reduction modelling results with shrinkaged exposures are stored. That list include:

1. "MedtRedXM_all": a dataframe containing the mediator dimension reduction result where all mediators are considered as one group and that shrinkaged mediator is paired and modelled with shrinkaged exposures.
2. "MedtRedXM_list": a dataframe containing the mediator dimension reduction result where mediators are shrinkaged in their own subgroups, and these shrinkaged mediators are paired and modelled with shrinkaged exposures.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
res1 <- LoadMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
VocaPath = NULL)
res2 <- XMlists(PID = res$PID)
res3 <- RedX(PID=res$PID, VarsY = "Y1",
VarsC = "default", VarsM = "default", Method = "mean", Folds = 10,
Family = "linear", Iter = 500)
res4 <- RedXM(PID=res$PID, VarsY = "Y1",
VarsC = "default", Method = "mean", Family = "linear", Iter = 500)
res5 <- RedXM(PID=res$PID, VarsY = "Y1",
VarsC = "C1", Method = "mean", Family = "linear", Iter = 500)
```

StatComp *Size comparison between groups*

Description

Size comparison between groups

Usage

```
StatComp(PID, OutPath="default", Group, Task = "mean",
   Vars, VarsBy, Method = "wilcox", Layout = "density", Brightness = "dark" ,
   Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStat.
OutPath	chr. Output file directory, e.g. "D:/ExpoStat/StatComp". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
Group	lgl. Whether to separate dataset into train and test data for normality test. The default is "TRUE".
Task	chr. Comparison task. At present, only the mean comparison is available.
Vars	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space. The default values is "all" (all variables are included).
VarsBy	chr. Variable used to group the observation for size comparison.
Method	chr. Comparison method. At present, only "wilcox" (Wilcoxon rank sum test) is available.
Layout	chr. Visualization layout. Available values include "column.points", "density".
Brightness	chr. Visualization brightness. Available values include "light" and "dark".
Palette	chr. Visualization palette. Available values include "default1", "default2" , "default3" and 5 journal option including "cell", "nature", "science", "lancet", "nejm".

Value

A list object containing the results of size comparisons between groups for variables and visualization of the results.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStat()
   res1 = LoadStat(PID = res$PID, UseExample = "example#1")
   res2 = StatComp(PID=res$PID, Group = T, Task = "mean", Vars = "X5,X6,X7,X8,X9",
   VarsBy = "Y1", Method = "wilcox", Layout = "density", Brightness = "dark",
   Palette = "default1")
```

StatCorr *Correlation analysis between variables*

Description

Correlation analysis between variables

Usage

```
StatCorr(PID, OutPath="default", Group, VarsX, VarsY, VarsBy,
  Method = "spearman", Layout= "bubble", Brightness = "dark" , Palette ="default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStat.
OutPath	chr. Output file directory, e.g. "D:/ExpoStat/StatCorr". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
Group	lgl. Whether to separate dataset into train and test data for normality test. The default is "TRUE".
VarsX	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space. The default values is "all.x" (all variables are included).
VarsY	chr. Target outcome variables used for correlation analysis.
VarsBy	chr. Variable used to group the observation for correlation analysis.
Method	chr. Method for orrelation analysis. Available values include "spearman" (Spearman's rank correlation analysis) and "pearson" (Pearson correlation analysis).
Layout	chr. Visualization layout. Available values include "heatmap", "bubble", "matrix".

Brightness	chr. Visualization brightness. Available values include "light" and "dark".
Palette	chr. Visualization palette. Available values include "default1", "default2" , "default3" and 5 journal option including "cell", "nature", "science", "lancet", "nejm".

Value

A list object containing the results of correlation analysis between variables and visualization of the results.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStat()
   res1 = LoadStat(PID = res$PID, UseExample = "example#1")
   res2 = StatCorr(PID = res$PID, Group = T, VarsX = "X5,X6,X7,X8,X9", VarsY = "Y1",
   VarsBy = "Y1", Method = "pearson", Layout = "bubble", Brightness = "dark",
   Palette = "nature")
```

StatDesc *Variable description*

Description

Variable description

Usage

```
StatDesc(PID, OutPath="default", Group, Vars, VarsBy, Layout="box",
   Brightness="light", Palette="default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStat.
OutPath	chr. Output file directory, e.g. "D:/ExpoStat/StatDesc". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
Group	lgl. Whether to separate dataset into train and test data for normality test. The default is "TRUE".
Vars	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space. The default values is "all" (all variables are included).
VarsBy	chr. Variable used to group the observation for size description.

Layout	chr. Visualization layout. Available values include "box", "violin".
Brightness	chr. Visualization brightness. Available values include "light" and "dark".
Palette	chr. Visualization palette. Available values include "default1", "default2" and 5 journal option including "cell", "nature", "science", "lancet", "nejm".

Value

A list object containing the results of variable description for continuous and discrete variables respectively and visualization of the continuous variables.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStat()
   res1 = LoadStat(PID = res$PID, UseExample = "example#1")
   res2 = StatDesc(PID = res$PID, Group = T, Vars = "C1,C2,X5,X6,X7,X8,X9",
   VarsBy = NULL, Layout = "box", Brightness = "dark", Palette = "default2")
```

StatExtre *Extreme value calculation*

Description

Extreme value calculation

Usage

```
StatExtre(PID, OutPath="default", Group, Vars, LimitLow = 0.025,
   LimitUpper = 0.975, Layout = "column.points", Brightness = "light",
   Palette = "default2")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStat.
OutPath	chr. Output file directory, e.g. "D:/ExpoStat/StatExtre". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
Group	lgl. Whether to separate dataset into train and test data for normality test. The default is "TRUE".
Vars	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space. The default values is "all" (all variables are included).
LimitLow	num. Lower limit ratio to screen the small extreme values located from 0 to this lower limit of the target variables.

LimitUpper	num. Upper limit ratio to screen the large extreme values located from this lower limit to 1 of the target variables.
Layout	chr. Visualization layout . Available values include "column.points", "heatmap".
Brightness	chr. Visualization brightness . Available values include "light" and "dark".
Palette	chr. Visualization palette . Available values include "default1", "default2", "default3" and 5 journal option including "cell", "nature", "science", "lancet", "nejm".

Value

A list object containing the results of extremum for variables and visualization of the results.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStat()
   res1 = LoadStat(PID = res$PID, UseExample = "example#1")
   res2 = StatExtre(PID = res$PID, Group = T, Vars = "X5,X6,X7,X8,X9",
   LimitLow = 0.025, LimitUpper = 0.975, Layout = "column.points",
   Brightness = "dark", Palette = "default2")
```

StatLinkCros *Build statistical link for cross-sectional data.*

Description

Build statistical link for cross-sectional data.

Usage

```
StatLinkCros(PID, OutPath = "default", VarsY, VarsX, LinkModel = "ranger",
    ObsrPartType = "raw",ObsrPartNum = "50",ObsrProfType = "partial",
    ObsrProfNum = "100",ObsrProfVars = "all.x",ObsrProfGeom = "profiles",
    SubjPredSeq = "none",SubjPartType = "break_down")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoStatLink
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modelling. Only one variable can be entered.

VarsX	chr. Exposure variable used for modelling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
LinkModel	chr. Methods to interpret the model. Options include "ranger" (random forest), "glmnet" (elastic net), "svm" (support vector machine), "glm" (linear regression), "gam" (generalized additive model), and "xgboost" (eXtreme gradient boosting).
ObsrPartType	chr. Type of transformation that should be applied for dropout loss. Options include "raw" (drop losses), "ratio" (drop_loss/drop_loss_full_model), and "difference" (drop_loss - drop_loss_full_model)
ObsrPartNum	chr. Number of observations that should be sampled for calculation of variable importance. The default means variable importance will be calculated on whole dataset (no sampling). If "defult", use all Obsrrvations.
ObsrProfType	chr. Type of variable profile. Options include "partial", "conditional", and "accumulated"
ObsrProfNum	int. Number of observations used for calculation of aggregated profiles. By default 100.
ObsrProfVars	chr. Names of variables to be explained. If "all.x", all "X variable" above are chosen.
ObsrProfGeom	chr. Layout of the explanation profile in dataset level including "aggregates", "profiles" or "points".
SubjPredSeq	chr. Subjects which need explanation. Options include "all" (all the subjects), "none" (no subjects), and "other" (copy the subject list by clicking "Available vars").
SubjPartType	chr. Layout of the explanation profile in subject level. Options include "shap", "oscillations", "break_down", and "none".

Value

A list containing all the statistical explanation results.

Author(s)

Bin Wang

Examples

```
res <- InitStatLink()
   res1 = LoadStatLink(PID = res$PID, UseExample = "example#1")
   res2 = StatLinkCros(PID=res$PID, VarsY = "Y1" ,VarsX = "all.x",
   LinkModel = "ranger",ObsrPartType = "raw" ,ObsrPartNum  = "50",
   ObsrProfType = "partial" ,ObsrProfNum = "100", ObsrProfVars = "all.x",
   ObsrProfGeom = "profiles",SubjPredSeq = "S1,S2,S3",SubjPartType = "break_down")
```

StatNorm *Normality test for numeric variables*

Description

Normality test for numeric variables

Usage

```
StatNorm(PID, OutPath="default", Group, Vars, Method = "shapiro.test", Layout = "rose.chart",
   Brightness = "light", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStat.
OutPath	chr. Output file directory, e.g. "D:/ExpoStat/StatNorm". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
Group	lgl. Whether to separate dataset into train and test data for normality test. The default is "TRUE".
Vars	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space. The default values is "all" (all variables are included).
Method	chr. Normality test method. Only "shapiro.test" method is available at present.
Layout	chr. Visualization layout. Available values include "column", "column.points", "rose.chart", and "density".
Brightness	chr. Visualization brightness. Available values include "light" and "dark".
Palette	chr. Visualization palette. Available values include "default1", "default2", "default3" and 5 journal option including "cell", "nature", "science", "lancet", "nejm".

Value

A list object containing the results of normality test for variables and visualization of the results.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStat()
   res1 = LoadStat(PID = res$PID, UseExample = "example#1")
   res2 = StatNorm(PID=res$PID, Group = T, Vars = 'X5,X6,X7,X8,X9',
   Method = "shapiro.test",Layout = "rose.chart" , Brightness = "dark",
   Palette = "default3")
```

StatTable1 *Create Table 1 for for different epidemilogical study designs*

Description

Create Table 1 for different epidemilogical study designs

Usage

```
StatTable1(PID, OutPath="default", EpiDesign = "cohort",
   Group, VarsY, VarsC, Missing = "ifany")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStat.
OutPath	chr. Output file directory, e.g. "D:/ExpoStat/StatTable1". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
EpiDesign	chr. Research types provided for users, include "cohort" , "case-control", "cross-section".
Group	lgl. Whether to separate dataset into train and test data for creating Table 1. The default is "TRUE".
VarsY	chr. Outcome variable used for modelling. Only one variable can be entered.
VarsC	chr. Covariate variables needing further statistical test. It should be noted that there is fixed format for the entering characters separated with "," and without space. The defaults value is all covariate variables listed in the data file, which can be entered with "all.c".
Missing	chr. Counts of missing values in the table, available options include are "no" (never display missing values), "ifany" (only display if any missing values), and "always" (includes missing count row for all variables). Default is "ifany".

Value

A list object containing standardized table 1.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStat()
   res1 = LoadStat(PID = res$PID, UseExample = "example#1")
   res2 = StatTable1(PID = res$PID, EpiDesign = "cohort" ,
   Group = 'T', VarsY = "Y1", VarsC = "C1,C2,C3,C4,C5,C6",
   Missing = "ifany")
```

SurvAsso *Association analysis*

Description

Association analysis for survival data

Usage

```
SurvAsso(PID, OutPath = "default",TimeY, EventY,VarsX='all.x',VarsN = "single.factor",VarsSel = T,
    IncCova = T)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TimeY	chr. Outcome variable of survival time used for modelling. Only one variable can be entered.
EventY	chr. Outcome variable of status used for modelling. Only one variable can be entered.
VarsX	Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
VarsSel	lgl. T (or TRUE) and F (or FALSE). Whether to select the significant variable for the final model. Available options includes T and F
IncCova	lgl. T (or TRUE) and F (or FALSE). Whether to include the covariate selected in the function "FindCovaSurv"

Value

A list containing the association analysis results.

Author(s)

Changxin Lan, Bin Wang(corresponding author)

Examples

```
res <- InitSurv()
   res1 = LoadSurv(PID = res$PID, UseExample = "example#1")
   res3 = FindCovaSurv(PID=res$PID, TimeY = "Y1", EventY= 'Y2',
   VarsC_Prior = "default", VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
   res4 = SurvAsso(PID=res$PID, TimeY = "Y1", EventY= 'Y2',VarsX='all.x',
   VarsN="single.factor",VarsSel=T,IncCova=T)
   FuncExit(PID = res$PID)
```

SurvPred *Build prediction models*

Description

Build prediction models

Usage

```
SurvPred(PID,OutPath = "default",TimeY,EventY,VarsX = "all.x",
    IncCova = T,RsmpMethod = "cv",Folds = 3,Ratio = 0.667,Repeats = 3)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TimeY	chr. Outcome variable of survival time used for modelling. Only one variable can be entered.
EventY	chr. Outcome variable of status used for modelling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
IncCova	lgl. Whether to include the covariate selected in the function of "FindCovaSurv". Available options include T (or TRUE) and F (or FALSE).
RsmpMethod	chr. Three resampling methods options for internal validation, including "cv" (i.e.,Cross validation) , "bootstrap", and "holdout".
Folds	num. Folds of Cross-validation resampling. It is ranging 2-10.
Ratio	num. Ratio of Bootstrap resampling. It is ranging 0.4-0.9.
Repeats	num. Number of Bootstrap resampling. It is ranging 2-20.

Value

A list containing the prediction performance evaluation.

Author(s)

Changxin Lan, Bin Wang(corresponding author)

Examples

```
res <- InitSurv()
   res1 = LoadSurv(PID = res$PID, UseExample = "example#1")
   res3 = FindCovaSurv(PID=res$PID, TimeY = "Y1", EventY= 'Y2',
   VarsC_Prior = "default", VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
   res4 = SurvPred(PID=res$PID, TimeY = "Y1", EventY= 'Y2',VarsX='all.x',
   IncCova=T,RsmpMethod="cv",Folds=3,Ratio=0.667,Repeats=3)
   FuncExit(PID = res$PID)
```

TransClass *Classify variables into various groups*

Description

Classify variables into various groups

Usage

```
TransClass(PID, OutPath = "default", Group, Vars, LevelTo)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Group	lgl. Whether to separate dataset into train and test data for processing data.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; "all.c", all covariates; "all.cx", combination of All X and All C. Users can also choose available variables by selecting "Other" option, and copy the variables by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
LevelTo	The number of levels to convert variables to.

Value

An R6 class object containing the variable(s) after classifying data into various levels.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
    res1 = LoadTidy(PID=res$PID, UseExample="example#1")
    res2 = TransClass(PID=res$PID, Group= FALSE, Vars="X1", LevelTo="4")
    FuncExit(PID = res$PID)
```

TransDistr *Transform variable distribution*

Description

Transform variable distribution

Usage

```
TransDistr(PID,OutPath = "default", Vars, Method)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; "all.c", all covariates; "all.cx", combination of All X and All C. Users can also choose available variables by selecting "Other" option, and copy the variables by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
Method	chr. Methods used for imputation. Available options include "lod" or "cart". For "lod" method, limit of detection (LOD) should be included in the "Vocabulary" file.

Value

An R6 class object containing the variable(s) after transforming distribution.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
    res1 = LoadTidy(PID=res$PID, UseExample="example#1")
    res2 = TransDistr(PID=res$PID, Vars="X6,X7", Method="log10")
    FuncExit(PID = res$PID)
```

TransDummy *Transform factor variables into dummy ones*

Description

Transform factor variables into dummy ones

Usage

```
TransDummy(PID, OutPath = "default", Vars="default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	chr. Variables to be transformed as dummy variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3". If "default", all the factor variables will be transformed into dummy ones. These variables need to be transformed as factor ones in previous transform step using TransType function.

Value

An R6 class object containing the variable(s) after transforming the factor variables into dummy ones.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
    res1 = LoadTidy(PID=res$PID, UseExample="example#1")
    res2 = TransDummy(PID=res$PID, Vars="default")
    FuncExit(PID = res$PID)
```

TransGroup *Transform exposure groups*

Description

Transform exposure groups

Usage

```
TransGroup(PID, OutPath = "default", Vars="default", ToGroup)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; "all.c", all covariates; "all.cx", combination of All X and All C. Users can also choose available variables by selecting "Other" option, and copy the variables by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
ToGroup	chr. Label the group of the target variables. Four common used names are recommended, including Exposure, Metabolome, Proteome, and Immunome. Users can also label the groups as you like.

Value

An R6 class object containing the variable(s) after grouping the variables.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
    res1 = LoadTidy(PID=res$PID, UseExample="example#1")
    res2 = TransGroup(PID=res$PID, Vars="X4,X5", ToGroup = "G1")
    FuncExit(PID = res$PID)
```

TransImput *Missing data imputation.*

Description

Missing data imputation.

Usage

```
TransImput(PID,OutPath = "default",Group,Vars,Method)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Group	lgl. Whether to separate dataset into train and test data for processing data.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; "all.c", all covariates; "all.cx", combination of All X and All C. Users can also choose available variables by selecting "Other" option, and copy the variables by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
Method	Methods used for imputation. Available options include "lod" or "cart" methods. For "lod" method, limit of detection (LOD) should be included in the "Vocabulary" file.

Value

An R6 class object containing variable(s) with imputation.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
    res1 = LoadTidy(PID=res$PID, UseExample="example#1")
    res2 = TransImput(PID=res$PID, Group=TRUE, Vars="all.x", Method="lod")
    FuncExit(PID = res$PID)
```

TransScale *Scale variables*

Description

Scale variables

Usage

```
TransScale(PID,OutPath = "default", Group = T, Vars, Method = "normal",
    Direct="positive", RangeLow="0", RangeUpper="1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Group	lgl. T (or TRUE) and F (or FALSE). Whether to separate dataset into train and test data for processing data.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; "all.c", all covariates; "all.cx", combination of All X and All C. Users can also choose available variables by selecting "Other" option, and copy the variables by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
Method	chr. Scaling methods. Available options include "normal" and "range".
Direct	chr. Direction to be transformed, Available options include "positive" and "negative".
RangeLow	num. Lower limit for range method.
RangeUpper	num. Upper limit for range method. It should be greater than the lower limit.

Value

An R6 class object containing the variable(s) after scaling data.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
    res1 = LoadTidy(PID=res$PID, UseExample="example#1")
    res2 = TransScale(PID=res$PID, Group= TRUE, Vars="all.x", Method="normal")
    FuncExit(PID = res$PID)
```

TransType *Transform data type*

Description

Transform data type

Usage

```
TransType(PID,OutPath = "default",Vars,To)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; "all.c", all covariates; "all.cx", combination of All X and All C. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
To	chr. Indicate the type of the chosen variables to be transformed into. Available options include "integer", "numeric", "character", "factor", "logical", and "date".

Value

An R6 class object containing the variable(s) after transforming data type.

Author(s)

Bin Wang

Examples

```
res = InitTidy()
    res1 = LoadTidy(PID=res$PID, UseExample="example#1")
    res2 = TransType(PID=res$PID, Vars = "X1,X2", To = "character")
    FuncExit(PID = res$PID)
```

VizBioLink *Visualize the biological link*

Description

Visualize the biological link. It should be noted that the corresponding link has been built by BioLink function prior to using it

Usage

```
VizBioLink(PID, OutPath="default", Mode, Layout = "force-directed", Brightness = "dark", Palette = "def
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitBioLink.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Mode	chr. Method to build the biological link between exposures and diseases. Available options include "PPI" (i.e., protein-protein interaction) and "GO" (i.e., gene ontoloty).
Layout	chr. Visualization layout. Available options include "force-directed" and "degree-circle".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list object containing the plot of the biological link. This plot can be further processed using ggplot2 package.

Author(s)

Mingliang Fang, Ning Gao, Bin Wang (corresponding author)

Examples

```
res = InitBioLink()
   res1 = LoadBioLink(PID = res$PID,UseExample = "example#1")
   res2 = ConvToExpoID(PID = res$PID)
   res3 = BioLink(PID = res$PID, OutPath="default", Mode = "PPI", ChemCas = "default",
   ChemInchikey = "default",DiseaseID = "default",MetabolomeID = "default",
   MetBiospec = "blood", ProteomeID = "default")
   res4 = VizBioLink(PID = res$PID, Mode = 'PPI', Layout = "force-directed",
   Brightness = "dark", Palette = "default1")
```

VizCateDot *Plot category dot*

Description

Visualize data via dot plot.

Usage

```
VizCateDot(PID,OutPath="default",Group="F",Vars,Parameter="mean",
    Brightness="light",Palette="default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitViz.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
Group	lgl.Whether to separate dataset into train and test datasets for data imputation,including T or F.The default option is F.
Vars	chr. Specifying the variables. Available options include: "all.x", all independent variables; "all.c", all covariate variables; "all.cx", combination of All x and All x; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10"). No more than 50 variables be entered is recommended (< 50 variables).
Parameter	chr. Specifying which parameter of the data to be the ordinate of the output plot. Available options include: "mean", "median", "min", "max", "mad" or "sd".Default is "mean".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

The dot plot is used to display the relative position of two data points in the same time period, or compare the difference between the two categorical variables.

Value

One or three ggplot plots which can be further modified using the ggplot2 package. When Group=T, (1) "all_light_default1": the whole dataset visualization result; (2) "train_light_default1": the train dataset visualization result; (3) "test_light_default1": the test dataset visualization result. When Group=F, (1) "light_default1": the whole dataset visualization result.

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitViz()
    res1 = LoadViz(PID = res$PID, UseExample = "example#1")
  res2 = VizCateDot(PID=res$PID,OutPath="default",Group="F",Vars="X4,X5,X6,X7,X8,X9,X10",
   Parameter="mean",Brightness="light",Palette="default1")
```

VizCompoDendrogram *Plot component dendrogram*

Description

Visualize data via dendrogram plot.

Usage

```
VizCompoDendrogram(PID,OutPath="default",Group = "T",Vars,
    Parameter = "median",DistMethod = "euclidean", ClusterMethod = "ward.D",
    ClusterNum = "4",Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitViz.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Group	lgl.Whether to separate dataset into train and test datasets for data imputation,including T or F.The default option is T.
Vars	chr. Variables to be visualized(e.g."X4,X5,X6,X7,X8,X9,X10").Available options include: "all.x", all independent variables; "all.c", all covariate variables; "all.cx", combination of all.x and all.c; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10").
Parameter	chr. Specifying which parameter of the data to be the ordinate of the output plot. Available options include: "mean", "median", "min", "max", "mad" or "sd".Default is "mean".
DistMethod	chr.The distance measure. This must be one of "euclidean", "maximum" or "manhattan".Default is "euclidean".
ClusterMethod	chr.The agglomeration method. This should be one of "ward.D", "ward.D2" or "single".Default is "ward.D".
ClusterNum	num. The number of groups for cutting the tree.Default is 4.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

The dendrogram plot is used to plot beautiful dendrograms.

Value

One or three ggplot plots which can be further modified using the ggplot2 package. When Group=T, (1) "all_light_default1": the whole dataset visualization result; (2) "train_light_default1": the train dataset visualization result; (3) "test_light_default1": the test dataset visualization result. When Group=F, (1) "light_default1": the whole dataset visualization result.

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitViz()
    res1 = LoadViz(PID = res$PID, UseExample = "example#1")
    res2= VizCompoDendrogram(PID=res$PID,OutPath="default",Group = "T",
    Vars = "X4,X5,X6,X7,X8,X61,X66,X67,X200",Parameter = "median",
    DistMethod = "euclidean",ClusterMethod = "ward.D2",ClusterNum = "4",
    Brightness = "light",Palette = "default1")
```

VizCrosAsso *Visualize association analysis*

Description

Visualize association analysis

Usage

```
VizCrosAsso(PID, OutPath = "default", VarsY, VarsN,
    Layout = "volcano", Brightness = "light", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoCros
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
Layout	chr. Visualization layout. Available options include "forest" and "volcano".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

An R6 class object containing the results' plot.

Author(s)

Bin Wang

Examples

```
res <- InitCros()
   res1 = LoadCros(PID = res$PID, UseExample = "example#1")
   res2 = CrosAsso(PID=res$PID, EpiDesign = "cohort", VarsY = "Y1",
   VarsX = "X5,X6,X7,X8,X9,X10,X11", VarsN = "single.factor",
   VarsSel = FALSE, VarsSelThr = 0.1, IncCova = TRUE, Family = "gaussian",
   RepMsr = FALSE,Corstr = "ar1")
   res3 = VizCrosAsso(PID=res$PID,VarsY = "Y1",VarsN="single.factor", Layout = "forest",
   Brightness = "dark",Palette = "default1")
   FuncExit(PID = res$PID)
```

VizCrosPred *Visualize the prediction performance*

Description

Visualize the prediction performance

Usage

```
VizCrosPred(PID, OutPath = "default", VarsY, Layout = "bar", Brightness = "light",
    Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoCros
OutPath	chr. Output file directory, e.g., "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
Layout	chr. Visualization layout. Available options include "bar" and "roc"
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the results' plot.

Author(s)

Bin Wang

Examples

```
res <- InitCros()
   res1 = LoadCros(PID = res$PID, UseExample = "example#1")
   res2 = CrosPred(PID=res$PID, VarsY = "Y1", VarsX = "X5,X6,X7,X8,X9,X10,X11",
   PredType = "response", VarsSel = FALSE, VarsSelThr = 0.1,IncCova = FALSE,
   RsmpMethod = "cv", Folds = 5, Ratio = 0.667, Repeats = 5)
   res3 = VizCrosPred(PID=res$PID, VarsY = "Y1", Layout = "bar", Brightness = "light",
   Palette = "science")
   FuncExit(PID = res$PID)
```

VizDistrSierra *Plot distribution sierra*

Description

Visualize data via sierra plot.

Usage

```
VizDistrSierra(PID,OutPath="default",Group = "F",Vars,Brightness = "light",
    Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitViz.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Group	lgl.Whether to separate dataset into train and test datasets for data imputation,including T or F.The default option is F.
Vars	chr. Variables to be visualized(e.g."X4,X5,X6,X7,X8,X9,X10").Available options include: "all.x", all independent variables; "all.c", all covariate variables; "all.cx", combination of all.x and all.c; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10"). No more than 50 variables be entered is recommended (< 20 variables).
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

The sierra plot is used to visualize the kernel density estimation of data.

Value

One or three ggplot plots which can be further modified using the ggplot2 package. When Group=T, (1) "all_light_default1": the whole dataset visualization result; (2) "train_light_default1": the train dataset visualization result; (3) "test_light_default1": the test dataset visualization result. When Group=F, (1) "light_default1": the whole dataset visualization result.

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitViz()
    res1 = LoadViz(PID = res$PID, UseExample = "example#1")
    res2 = VizDistrSierra(PID=res$PID,OutPath="default",Group = "F",
    Vars = "X14,X15,X16,X17,X18,X19,X20",Brightness = "light",Palette = "default1")
```

VizMedtPair *Visualize pairwise mediation modelling result*

Description

Visualize the pairwise mediation result. It should be noted that the pairwise mediation modelling result has been built by Pairwise function prior to using it.

Usage

```
VizMedtPair(PID, OutPath, Brightness = "light", Pallette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Pallette	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

VizMedtPair draws a visualized display for pairwise modelling results. Piror to using VizMedtPair, make sure that the users have built mediation models by Pairwise function.

Value

A list containing two elements where the visualized pairwise modelling plot as well as the organized plotting data are stored. The elements of that list include:

1. "plotdata": a dataframe containing the organized plotting data extracted from #' pairwise mediation modelling result.
2. "plot": a visualized plot for pairwise mediation result.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
res1 <- LoadMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
VocaPath = NULL)
res2 <- XMlists(PID = res$PID)
res3 <- PairWise(PID=res$PID, VarsY = "Y1",
VarsX = "default", VarsM = "default", VarsC = "default", Family = "linear",
Iter = 500)
res4 <- VizMedtPair(PID=res$PID, Brightness = "Bright",
Pallette = "default1")
```

VizMixBKMR *Visualize the model results of Bayesian Kernel Machine Regression (BKMR)*

Description

Visualize the model results of Bayesian Kernel Machine Regression (BKMR)

Usage

```
VizMixBKMR(PID, OutPath = "default", VarsY, Brightness = "dark", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoMixEffect
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the BKMR analysis results' plot.

Author(s)

Bin Wang

Examples

```
res <- InitMix()
   res1 = LoadMix(PID = res$PID, UseExample = "example#1")
   res2 = MixBKMR(PID = res$PID, VarsY = "Y1", VarsX = "X4,X5,X6,X7,X8,X9,X10",
   IncCova = 'F', Family = "gaussian", Group = 'F',Iter = 2000,qfixed = 0.5,
   qsbivar = "default", qsoverall = "default",qsdiff = "default")
   res3 = VizMixBKMR(PID=res$PID, VarsY = "Y1",Brightness = "dark",Palette = "default1")
```

VizMixMLR *Visualize the model results*

Description

Visualize the results of multiple linear regression (MLR) model

Usage

```
VizMixMLR(PID,OutPath = "default",VarsY,SelMethod,Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoMixEffect
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
SelMethod	chr. Method to select the important features to the final model. Options include "stepwise" (stepwise regression), "lasso" (Regularization regression of least absolute shrinkage and selection operator), and "enet" (Regularization regression of elastic net).
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the MLR analysis plot.

Author(s)

Bin Wang

Examples

```
res <- InitMix()
   res1 = LoadMix(PID = res$PID, UseExample = "example#1")
   res2 = MixMLR(PID = res$PID, VarsY = "Y1", VarsX = "all.x",
   IncCova = "F", SelMethod = "lasso", PredType = "response", Family = "gaussian")
   res3 = VizMixMLR(PID=res$PID, VarsY = "Y1", SelMethod = 'lasso',
   Brightness = "light", Palette = "default1")
```

VizMixWQS *Visualize results of weighted quantile sum regression (WQS) model*

Description

Visualize results of weighted quantile sum regression (WQS) model

Usage

```
VizMixWQS(PID, OutPath = "default", VarsY, Brightness = "dark", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoMixEffect
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the WQS analysis results' plot.

Author(s)

Bin Wang

Examples

```
res <- InitMix()
   res1 = LoadMix(PID = res$PID, UseExample = "example#1")
   res2 = MixWQS(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   IncCova = "F",Family = "gaussian", VarStrat = "none", RatioValidat = 0.3,
   q = 10, b=100, b1_pos = 'F', b1_constr = 'F')
   res3 = VizMixWQS(PID = res$PID,  VarsY = "Y1",
   Brightness = "dark",Palette = "default1")
```

VizMulOmicCros *Visualize multiomics model results*

Description

VizMulOmicCros function is mainly aimed to visualize the modeling results calculated by MulOmicsCros function. It can provide plots with high quality of the final results to make it easier for users to understand.

Usage

```
VizMulOmicCros(PID, OutPath, VarsY, NodeNum, EdgeThr, Layout, Brightness, Palette)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMo.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable for modeling. Only one variable can be entered.
NodeNum	num. Number of nodes in the network plot. The maximum number is generated by Multiomics function. User can set a smaller value if needed.
EdgeThr	num. Threshold of correlation coefficient ranging 0-1 for generating the concerned edges of the network plot.
Layout	chr. Visualization layout. Available options include "force-directed" and "degree-circle".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and other options about some journal preference styles including "cell", "nature", "science", "lancet", "nejm", etc.

Details

You can get different styles of images by selecting different parameters.

Value

An R6 class object containing seven elements. The elements of that object include: (1) "Importance_plot": Plots for the importance of features after modeling a single omic from different learners. (2) "Measures_boxplot": Plots for the r-square value of the single omic model built by different learners. (3) "NetWork_State": A list containing dataframes that contain nodes and edges used to draw interOmic network from models built by different learners. (4) "Nodeplot": Interomic node plots from models built by different models. (5) "Networkplot": Interomic network plots from models built by different models. (6) "Prediction_train_plot": Plots for the prediction value of the train set in the SG model built by different combinations of learners. (7) "Prediction_test_plot": Plots for the prediction value of the test set in the SG model built by different combinations of learners.

Author(s)

Guohuan Zhang, Yuting Wang, Ning Gao, Bin Wang (corresponding author)

Examples

```
 res <- InitMO()
 res <- LoadMO(PID=res$PID, UseExample = "example#1", DataPath = NULL, VocaPath = NULL)
res2 <- MulOmicsCros(PID=res$PID, OutPath = "default", OmicGroups = "immunome,metabolome,proteome",
VarsY = "Y1", VarsC = "all.c", TuneMethod = "random_search", TuneNum = 5, RsmpMethod = "cv", Folds = 5,
 Ratio = 0.67, Repeats = 5, VarsImpThr = 0.85, SG_Lrns ="lasso,enet,rf,xgboost")
res3 <- VizMulOmicCros(PID=res$PID, OutPath = "default", VarsY = "Y1", NodeNum=100, EdgeThr= 0.45,
 Layout = "force-directed", Brightness = "light", Palette = 'default1')
```

VizNtaAnno *Visualize annotation results*

Description

Visualize annotation results of non-targeted data

Usage

```
VizNtaAnno(PID, OutPath = "default", VarsY, VarsN ,Accuracy = 1,
    Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
Accuracy	num. Upper limit of accuracy to match the target molecular. The default is 1.

Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing plots of non-target annotation analysis results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = NtaCros(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = T, SelMethod = "all",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, IncCova = "F", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
   res3 = NtaAnno(PID=res$PID,VarsY = "Y1",VarsX = "default",VarsN = "single.factor",
   FdrCorrect = "F",AdductPos = "M+H",AdductNeg = "M-H",Accuracy = 1)
   res4 = VizNtaAnno(PID=res$PID, VarsY = "Y1",VarsN = "single.factor",
   Accuracy = 1,Brightness = "light",Palette = "default1")
```

VizNtaCros *Visualize the results of Association analysis*

Description

Visualize the results of Association analysis for non-targeted data

Usage

```
VizNtaCros(PID, OutPath = "default", VarsY,VarsN, Layout = "volcano",
    EffectThr = 0.5, Brightness = "light", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
Layout	chr. Visualization layout. Available options include "forest" and "volcano".

EffectThr	num. Threshold of the total importance for the variables to a random forest model. It ranges 0.5-1.0 with the default value of 0.9.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing plots of non-target analysis association results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = NtaCros(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = T, SelMethod = "all",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, IncCova = "F", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
   res3 = VizNtaCros(PID=res$PID, VarsY = "Y1",VarsN = "single.factor",
   Layout = "volcano", EffectThr = 0.5, Brightness = "light",Palette = "default1")
```

VizPanelAsso *Visualize the results of association analysis for panel data*

Description

Visualize the results of association analysis for panel data

Usage

```
VizPanelAsso(PID, OutPath = "default", VarsY, VarsN, EffectThr = 0.5,
    Layout = "volcano", Brightness = "dark", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoPanel
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
EffectThr	num. Insert the cutoff line for the effect values.

Layout	chr. Visualization layout. Available options include "forest" and "volcano".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the plots of association analysis results.

Author(s)

Bin Wang

Examples

```
res <- InitPanel()
   res1 = LoadPanel(PID = res$PID, UseExample = "example#1")
   res2 = PanelAsso(PID=res$PID, VarsY = "Y1",
   VarsX = "X1,X2,X3,X4,X5,X6,X7,X8,X9,X10,X11,X12", VarsN = "single.factor",
    VarsRandomIpt = "SubjectID", VarsRandomSlp = "none", IncCova = F)
   res3 = VizPanelAsso(PID = res$PID, VarsY = "Y1",
    VarsN = "single.factor", Layout = "forest", Brightness = "dark",Palette = "default1")
```

VizRedXM *Visualize RedXM mediation modelling result*

Description

Visualize the RedXM mediation result. It should be noted that the RedXM mediation modelling result has been built by RedXM function prior to using it.

Usage

```
VizRedXM(PID, OutPath, Brightness = "light", Pallette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
Brightness	chr. Visualization brightness. Available options include "Light" and "Dark".
Pallette	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

VizRedXM draws a visualized display for RedXM results. Piror to using VizMedtPair, make sure that the users have built mediation models by RedXM function.

Value

A list containing two elements where the visualized exposure and mediator dimension reduction plot as well as the organized plotting data are stored. The elements of that list include:

1. "MedtRedXM_plotadta": a dataframe containing the organized plotting data extracted from exposure and mediator dimension reduction result.
2. "MedtRedXM_plot": a visualized plot for exposure and mediator dimension reduction result.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
res1 <- LoadMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
VocaPath = NULL)
res2 <- XMlists(PID = res$PID)
res3 <- RedX(PID=res$PID, VarsY = "Y1",
VarsC = "default", VarsM = "default", Method = "mean", Folds = 10,
Family = "linear", Iter = 500)
res4 <- RedXM(PID=res$PID, VarsY = "Y1",
VarsC = "default", Method = "mean", Family = "linear", Iter = 500)
res5 <- VizRedXM(PID=res$PID, Brightness = "bright",
Pallette = "nature")
```

VizRefer *Visualize the Articles' Main Information*

Description

Visualize the articles' main information after MetaRefer function.

Usage

```
VizRefer(PID, OutPath = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMeta.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set. It should be noted that the slash symbol is "/", not "\".

Details

VizRefer provides the visual function of articles' main information, showing the year and region distribution directly. Please attention, PID must be got from the return result of InitMeta(). VizRefer can only run successfully after successfully running InitMeta, LoadMeta and MetaRefer functions.

Value

A list object containing plots of article information visualization.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitMeta()
   res1 <- LoadMeta(PID = res$PID,
   UseExample = "example#1",
   DataPath = NULL)
   res2 <- MetaRefer(PID = res$PID,  OutPath = "default", Mode = "search",
   VarX = "default", VarY = "default", VarM = "default", YearFrom = "default",
   YearEnd = "default", PMID = "default")
   res3 <- VizRefer(PID=res$PID, OutPath = "default")
```

VizRelatEdgeBundling *Plot elationship edge bundling*

Description

Visualize data via edge bundling plot.

Usage

```
VizRelatEdgeBundling(PID,OutPath="default",VarsY, VarsC="all.c", VarsX = "all.x",
    Family = "gaussian", SizeFor = "pvalue",Brightness = "light",
    Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitViz.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
VarsY	chr. Dependent variables for visualization(e.g."Y2").
VarsC	chr. Covariate variable.Available options include: "all.c", all covariate variables; or input a character string specifying the variables,separated by comma "," without space(e.g."C1,C2").

VarsX	chr. Independent variables.Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10").
Family	chr. The link function for the regression model according the data type of outcomes,including "gaussian" for continuous variable, "binomial" for binary variable,and "poisson" for counting variable. Notice that the family are determined by data type of an outcome, or the plot can not be visualized.The default option is "gaussian".
SizeFor	chr. Parameter to represent the size of the points in the output plot. Available options include "pvalue" and "beta".The default option is "pvalue".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

The edge bundling plot is used to bundle the edges closely in order to reduce complexity.

Value

A ggplot plot which can be further modified using the ggplot2 package. (1) "light_default1": the visualization result;

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitViz()
    res1 = LoadViz(PID = res$PID, UseExample = "example#1")
    res2 = VizRelatEdgeBundling(PID=res$PID,OutPath="default",VarsY = "Y2",
    VarsX = "all.x",VarsC = "all.c",Family = "gaussian" ,SizeFor = "pvalue",
    Brightness = "light",Palette = "default1")
```

VizRelatHeatmap *Plot relationship heatmep*

Description

Visualize data via heatmap plot.

Usage

```
VizRelatHeatmap(PID,OutPath="default",Group = "F",VarsY,VarsX,
    Method = "spearman",Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitViz.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Group	lgl.Whether to separate dataset into train and test datasets for data imputation,including T or F.The default option is F.
VarsY	chr. Dependent variables for visualization(e.g."Y2").
VarsX	chr. Independent variables.Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10").
Method	chr. Method to calculate the correlation. Default option is "spearman".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

The heatmap plot is used to display data in color changes as a matrix.

Value

One or two ggplot plots which can be further modified using the ggplot2 package. When Group=T, (1) "train_light_default1": the train dataset visualization result; (2) "test_light_default1": the test dataset visualization result. When Group=F, (1) "light_default1": the whole dataset visualization result.

Author(s)

Ning Gao, Bin Wang (corresponding author)

Examples

```
res = InitViz()
    res1 = LoadViz(PID = res$PID, UseExample = "example#1")
    res2 = VizRelatHeatmap(PID=res$PID,OutPath="default",Group = "F",
    VarsY = "Y2",VarsX = "X1,X4,X5,X6,X7,X8,X9,X10",Method = "spearman",
    Brightness = "light",Palette = "default1")
```

VizRelatMatrix *Plot relationship matrix*

Description

Visualize data via matrix plot.

Usage

```
VizRelatMatrix(PID,OutPath="default",Group = "F",VarsY,VarsX,Method = "spearman")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitViz.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted t hat the slash symbol is "/", not "\". If "default", the current working directory will be set.
Group	lgl.Whether to separate dataset into train and test datasets for data imputation,including T or F.The default option is F.
VarsY	chr. Dependent variables for visualization(e.g."Y2").
VarsX	chr. Independent variables.Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10"). Notice that no more than 20 variables can be inputed.(< 20 variables).
Method	chr. Method to calculate the correlation. Default option is "spearman".

Details

The matrix plot is used to make a matrix of plots with a given data set.

Value

One or two ggmatrix plots which can be further modified using the ggplot2 package. When Group=T, (1) "Train": the train dataset visualization result; (2) "Test": the test dataset visualization result. When Group=F, (1) "All": the whole dataset visualization result.

Author(s)

Ning Gao, Bin Wang (corresponding author)

Examples

```
res = InitViz()
    res1 = LoadViz(PID = res$PID, UseExample = "example#1")
    res2 = VizRelatMatrix(PID=res$PID,OutPath="default",Group = "F",
    VarsY = "Y2",VarsX = "X4,X5,X6,X7,X8,X9,X10",Method = "spearman")
```

`VizRelatNetwork` *Plot relationship network*

Description

Visualize data via network plot.

Usage

```
VizRelatNetwork(PID,OutPath="default",VarsY,VarsC="all.c",VarsX="all.x",
    Family="gaussian",Layout="force-directed",CutOff=0.8,Brightness="light",
    Palette="default1")
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitViz.
`OutPath`	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
`VarsY`	chr. Specifying the dependent variables(e.g."Y2").
`VarsC`	chr. Specifying the covariate variable.Available options include: "all.c", all covariate variables; or input a character string specifying the variables,separated by comma "," without space(e.g."C1,C2").
`VarsX`	chr. Specifying the independent variables.Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10").
`Family`	chr. The link function for the regression model according the data type of outcomes,including "gaussian" for continuous variable, "binomial" for binary variable,and "poisson" for counting variable. Notice that the family are determined by data type of an outcome, or the plot can not be visualized.The default option is "gaussian".
`Layout`	chr. Visualization layout. Available options include "force-directed" and "degree-circle".
`CutOff`	num. Partial outcomes to visualize which is determined by correlation coefficient r. The range must between 0 and 1.
`Brightness`	chr. Visualization brightness. Available options include "light" and "dark".
`Palette`	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

The network plot is used to visualize the relationship between the input variables by using ggraph package.

Value

A ggplot plot which can be further modified using the ggplot2 package. (1) "light_default1": the visualization result;

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitViz()
    res1 = LoadViz(PID = res$PID, UseExample = "example#1")
    res2 = VizRelatNetwork(PID=res$PID,OutPath="default",VarsY="Y2",VarsC="all.c",
    VarsX="all.x",Family="gaussian",Layout="force-directed",CutOff=0.8,
    Brightness="light",Palette="default1")
```

VizSurvAsso *Visualize association analysis*

Description

Visualize association analysis

Usage

```
VizSurvAsso(PID, OutPath = "default",VarsN = "single.factor",
   Layout = "volcano",Brightness = "light",Palette = "default1",ColorFor= "p.value",SizeFor= "p.value"
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
Layout	chr. Visualization layout. Available options include "forest" and "volcano".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., "cell", "nature", "science", "lancet", "nejm", and "jama").
ColorFor	chr. Volcano plot dot color. Available options include "p.value" and "hr".
SizeFor	chr. Volcano plot dot size. Available options include "p.value" and "hr".

Value

A list containing the results' plot.

Author(s)

Changxin Lan, Ning Gao, Bin Wang(corresponding author)

Examples

```
res <- InitSurv()
   res1 = LoadSurv(PID = res$PID, UseExample = "example#1")
   res3 = FindCovaSurv(PID=res$PID, TimeY = "Y1", EventY= 'Y2',
   VarsC_Prior = "default", VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
   res4 = SurvAsso(PID=res$PID, TimeY = "Y1", EventY= 'Y2',VarsX='all.x',
   VarsN="single.factor",VarsSel=T,IncCova=T)
  res5 = VizSurvAsso(PID=res$PID,VarsN="single.factor",Layout="volcano",Brightness= "light",
   Palette = "default1",ColorFor= "p.value",SizeFor= "p.value")
   FuncExit(PID = res$PID)
```

VizSurvCompGroup *Compare the survival curves of two groups*

Description

Compare the survival curves of two groups

Usage

```
VizSurvCompGroup(PID,OutPath = "default",TimeY,EventY,
   VarsG,Model='km',VarsAdj,AdjMethod='average',Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TimeY	chr. Outcome variable of survival time used for modelling. Only one variable can be entered.
EventY	chr. Outcome variable of status used for modelling. Only one variable can be entered.
VarsG	chr. Grouping variable, must be a binary variable.
Model	chr. Methods to depict the survival curve. Options include 'km' (Kaplan-Meier estimate) and "coxph" (Cox proportional hazards regression mode).
VarsAdj	If you choose the cox model, co-variables used for modelling. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., X1,X2,X3.
AdjMethod	If you choose the cox model, method for adjusting model, include: "average","single","margin" and "conditional".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".

Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., "cell", "nature", "science", "lancet", "nejm", and "jama").

Value

A list containing the results' plot.

Author(s)

Changxin Lan, Ning Gao, Bin Wang(corresponding author)

Examples

```
res <- InitSurv()
   res1 = LoadSurv(PID = res$PID, UseExample = "example#1")
   res6 = VizSurvCompGroup(PID=res$PID,TimeY="Y1",EventY="Y2",VarsG="C3",
   Model="km",Brightness="light",Palette='default1')
   FuncExit(PID = res$PID)
```

VizSurvPred *Visualize the prediction performance*

Description

Visualize the prediction performance

Usage

```
VizSurvPred(PID,OutPath = "default",Layout = "curve",
    Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Layout	chr. Visualization layout. Available options include "curve" , "bar" and 'all'.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., "cell", "nature", "science", "lancet", "nejm", and "jama").

Value

A list containing the results' plot.

Author(s)

Changxin Lan, Ning Gao, Bin Wang(corresponding author)

Examples

```
res <- InitSurv()
   res1 = LoadSurv(PID = res$PID, UseExample = "example#1")
   res3 = FindCovaSurv(PID=res$PID, TimeY = "Y1", EventY= 'Y2',
   VarsC_Prior = "default", VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
   res4 = SurvPred(PID=res$PID, TimeY = "Y1", EventY= 'Y2',VarsX='all.x',
   IncCova=T,RsmpMethod="cv",Folds=3,Ratio=0.667,Repeats=3)
   res5 = VizSurvPred(PID=res$PID,Layout="curve",Brightness="light",Palette='default1')
   FuncExit(PID = res$PID)
```

XMlists *Divide exposures and mediators into different groups*

Description

Divide exposures and mediators into different groups according to the information specified in vocabulary file.

Usage

```
XMlists(PID, OutPath)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".

Details

XMlists divides all exposures and mediators into different groups. Therefore, the group information(i.e., the "Subgroup" variable) in vocabulary file is essential for XMlists function. Before using XMlist, the users must provide the information in advance and upload it by LoadMedt function.

Value

A list containing two lists where exposure and mediator variables were respectively categorized into subgroups in the form of dataframe. The elements of that list include:

1. "ExpoList": a list containing several dataframes that represent various exposure subgroups.
2. "MediList": a list containing several dataframes that represent various mediator subgroups.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitMedt()
   res1 <- LoadMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 <- XMlists(PID = res$PID)
```

Index