Package: gsmr2 (via r-universe)

September 9, 2024

Maintainer Zhihong Zhu <z.zhu1@uq.edu.au>, Angli Xue
   <a.xue@garvan.org.au>, Jian Yang <jian.yang@westlake.edu.cn>

Author Zhihong Zhu, Angli Xue, Zhili Zheng, Futao Zhang, Jian Yang

License GPL (>= 2.0)

URL <https://yanglab.westlake.edu.cn/software/gsmr/>

Type Package

Title gsmr2 - a tool for causal inference between complex traits

Description GSMR2 (Generalised Summary-data-based Mendelian
   Randomisation v2) is an improved version of GSMR, which uses
   GWAS summary statistics to test for a putative causal
   association between two phenotypes (e.g., a modifiable risk
   factor and a disease) based on a multi-SNP model. This version
   implements a global heterogeneity test to remove invalid
   instrumental variables and provides a causal estimation that is
   more robust to directional pleiotropy.

Version 1.1.1

Date 2024-1-31

Depends R (>= 2.15), methods, utils, stats, markdown

Suggests knitr, rmarkdown

Imports survey

LazyData no

NeedsCompilation no

VignetteBuilder knitr

RoxygenNote 7.0.2

Repository https://mrcieu.r-universe.dev

RemoteUrl https://github.com/jianyanglab/gsmr2

RemoteRef HEAD

RemoteSha e7953cc136ad1cfae14674facde21e90caddea5f

Contents

---

gsmr-package   *gsmr: A tool for GSMR and HEIDI analysis*

Description

Perform Generalized Summary-data-based Mendelian Randomization analysis (GSMR) and HEterogeneity In Dependent Instruments analysis to remove pleiotropic outliers (HEIDI-outlier).

Author(s)

Zhihong Zhu <z.zhu1@uq.edu.au>

Zhili Zheng <zhili.zheng@uq.edu.au>

Futao Zhang <f.zhang5@uq.edu.au>

Jian Yang <jian.yang@uq.edu.au>

References

Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. Nature Communications, in press. An early verison of the manuscript is available at bioRxiv, 168674.

---

bi_gsmr   *Bi-directional GSMR analysis*

Description

Bi-directional GSMR analysis is composed of a forward-GSMR analysis and a reverse-GSMR analysis that uses SNPs associated with the disease (e.g. at < 5e-8) as the instruments to test for putative causal effect of the disease on the risk factor.

Usage

```
bi_gsmr(bzx, bzx_se, bzx_pval, bzy, bzy_se, bzy_pval, ldrho, snpid, heidi_outlier_flag=T, gwas_thresh=
```

Arguments

`bzx`	vector, SNP effects on risk factor
`bzx_se`	vector, standard errors of bzx
`bzx_pval`	vector, p values for bzx
`bzy`	vector, SNP effects on disease
`bzy_se`	vector, standard errors of bzy
`bzy_pval`	vector, p values for bzy
`ldrho`	LD correlation matrix of the SNPs
`snpid`	genetic instruments
`n_ref`	sample size of the reference sample
`heidi_outlier_flag`	
	flag for HEIDI-outlier analysis
`gwas_thresh`	threshold p-value to select instruments from GWAS for risk factor
`single_snp_heidi_thresh`	
	p-value threshold for single-SNP-based HEIDI-outlier analysis
`multi_snp_heidi_thresh`	
	p-value threshold for multi-SNP-based HEIDI-outlier analysis
`nsnps_thresh`	the minimum number of instruments required for the GSMR analysis (we do not recommend users to set this number smaller than 10)
`ld_r2_thresh`	LD r2 threshold to remove SNPs in high LD
`ld_fdr_thresh`	FDR threshold to remove the chance correlations between SNP instruments
`gsmr2_beta`	GSMR2 beta version, including a new HEIDI-outlier method (used in a GSMR analysis) that is currently under development and subject to future changes, 0 - the original HEIDI-outlier method, 1 - the new HEIDI-outlier method

Value

Estimate of causative effect of risk factor on disease (forward_bxy), the corresponding standard error (forward_bxy_se), p-value (forward_bxy_pval) and SNP index (forward_index), and estimate of causative effect of disease on risk factor (reverse_bxy), the corresponding standard error (reverse_bxy_se), p-value (reverse_bxy_pval), SNP index (reverse_index), SNPs with missing values, with non-significant p-values and those in LD.

Examples

```
data("gsmr")
gsmr_result = bi_gsmr(gsmr_data$bzx, gsmr_data$bzx_se, gsmr_data$bzx_pval, gsmr_data$bzy, gsmr_data$bzy_se, gsmr_
```

---

gsmr *Generalized Summary-data-based Mendelian Randomization analysis*

Description

GSMR (Generalised Summary-data-based Mendelian Randomisation) is a flexible and powerful approach that utilises multiple genetic instruments to test for causal association between a risk factor and disease using summary-level data from independent genome-wide association studies.

Usage

```
gsmr(bzx, bzx_se, bzx_pval, bzy, bzy_se, ldrho, snpid, heidi_outlier_flag=T, gwas_thresh=5e-8, single_
```

Arguments

bzx	vector, SNP effects on risk factor
bzx_se	vector, standard errors of bzx
bzx_pval	vector, p values for bzx
bzy	vector, SNP effects on disease
bzy_se	vector, standard errors of bzy
ldrho	LD correlation matrix of the SNPs
snpid	genetic instruments
n_ref	sample size of the reference sample
heidi_outlier_flag	
	flag for HEIDI-outlier analysis
gwas_thresh	threshold p-value to select instruments from GWAS for risk factor
nsnps_thresh	the minimum number of instruments required for the GSMR analysis (we do not recommend users to set this number smaller than 10)
ld_r2_thresh	LD r2 threshold to remove SNPs in high LD
ld_fdr_thresh	FDR threshold to remove the chance correlations between SNP instruments
gsmr2_beta	GSMR2 beta version, including a new HEIDI-outlier method (used in a GSMR analysis) that is currently under development and subject to future changes, 0 - the original HEIDI-outlier method, 1 - the new HEIDI-outlier method
single_heidi_thresh	
	p-value threshold for single-SNP-based HEIDI-outlier analysis
multi_heidi_thresh	
	p-value threshold for multi-SNP-based HEIDI-outlier analysis

Value

Estimate of causative effect of risk factor on disease (bxy), the corresponding standard error (bxy_se), p-value (bxy_pval), SNP index (used_index), SNPs with missing values, with non-significant p-values and those in LD.

Examples

```
data("gsmr")
gsmr_result = gsmr(gsmr_data$bzx, gsmr_data$bzx_se, gsmr_data$bzx_pval, gsmr_data$bzy, gsmr_data$bzy_se, ldrho, g
```

---

std_effect	*Standardization of effect size and its standard error*

Description

Standardization of SNP effect and its standard error using z-statistic, allele frequency and sample size

Usage

```
std_effect(snp_freq, b, se, n)
```

Arguments

snp_freq	vector, allele frequencies
b	vector, SNP effects on risk factor
se	vector, standard errors of b
n	vector, per-SNP sample sizes for GWAS of the risk factor

Value

Standardised effect (b) and standard error (se)

Examples

```
data("gsmr")
std_effects = std_effect(gsmr_data$a1_freq, gsmr_data$bzx, gsmr_data$bzx_se, gsmr_data$bzx_n)
```

Index