

# Integrative Mendelian randomization for detecting exposure-by-group interactions using group-specific and combined summary statistics

Ke Xu, Nathaniel Maydanchik, Bowei Kang, Jianhai Chen, Qixiang Chen, Gongyao Xu, Shinya Tasaki, David A. Bennett, Lin S. Chen

## Introduction

This vignette provides an introduction to the `int2MR` R package. The `int2MR` package implements the `int2MR` method for detecting group-specific exposure-outcome effects and exposure-group interaction effects using summary statistics.

Before installing `int2MR`, ensure that the `devtools` package is installed. `int2MR` also requires `rstan` for Bayesian modeling. To install and load the development version of `int2MR`, run:

```
# Load devtools package
suppressMessages(library(devtools))

# Install int2MR from GitHub
install_github("Likeli-Ke/int2MR")
```

```
## Skipping install of 'int2MR' from a github remote, the SHA1 (7f974c81) has not changed since last in
## Use 'force = TRUE' to force installation
```

```
# Load int2MR
library(int2MR)

# Load rstan
suppressPackageStartupMessages(library(rstan))
```

```
## Warning: package 'rstan' was built under R version 4.3.3
```

```
## Warning: package 'StanHeaders' was built under R version 4.3.3
```

## Input Data Format

The `int2MR` package supports two types of input:

- **Two-sample data:** For integrating two sets of IV-to-outcome GWAS summary statistics.
- **Three-sample data:** For Integrating three sets of IV-to-outcome GWAS summary statistics.

Two-sample input data must be a list with the following components: `p` is the number of IVs; `hat_gamma` and `hat_s_gamma_sq` are the estimated IV-to-exposure effects and their variances (squared standard errors); `hat_Gamma1` and `hat_s1_sq` are the estimated IV-to-outcome effects for the first IV-to-outcome GWAS summary statistic and their variances; `hat_Gamma2` and `hat_s2_sq` are the estimated IV-to-outcome effects for the second IV-to-outcome GWAS summary statistic and their variances; `rho1` is the proportion of the comparison group in the sample of the first GWAS summary statistic; and `rho2` is the proportion of the comparison group in the sample of the second GWAS summary statistic.

```
# Load the example two-sample data
data(example_2sample_data)
names(example_2sample_data)
```

```
## [1] "p"                "hat_s1_sq"        "hat_s2_sq"        "hat_gamma"
## [5] "hat_Gamma1"       "hat_Gamma2"       "hat_s_gamma_sq"   "rho1"
## [9] "rho2"
```

Three-sample input data additionally contains `hat_Gamma3`, `hat_s3_sq`, and `rho3` for the third GWAS summary statistic.

```
# Load the example three-sample data
data(example_3sample_data)
names(example_3sample_data)
```

```
## [1] "p"                "hat_s1_sq"        "hat_s2_sq"        "hat_s3_sq"
## [5] "hat_gamma"       "hat_Gamma1"       "hat_Gamma2"       "hat_Gamma3"
## [9] "hat_s_gamma_sq" "rho1"            "rho2"            "rho3"
```

To detect the exposure-group interaction effect, the proportion of the comparison group in the samples (`rho`) must vary among the provided GWAS summary statistics.

## Running Examples

This section demonstrates how to run `int2MR` using simulated data. Two examples are provided: one for two-sample data and one for three-sample data.

### Example 1: Two-Sample Data

In this example, we perform an analysis on simulated two-sample data. The parameters include:

- **data\_list\_2sample:** Input data for two-sample analysis. Note that the `data_list_3sample` argument is left null
- **model\_type:** “2sample”
- **Prior distributions:** Inverse gamma priors with a shape and scale of 0.1
- **MCMC Settings:** 2 chains, 10,000 iterations with a warm-up period of 2,500 iterations, and an `adapt_delta` of 0.95.

```
result_2sample <- int2MR(data_list_2sample = example_2sample_data,
  model_type = "2sample",
  prior_inv_gamma_shape = 0.1,
  prior_inv_gamma_scale = 0.1,
  chains = 2, iter = 10000, warmup = 2500,
  adapt_delta = 0.95)
```

```
# Display the results for the two-sample analysis
result_2sample$result_2sample
```

```
##   est_beta   se_beta   pval_beta est_beta_int se_beta_int pval_beta_int
## 1 0.1910384 0.03606093 1.172981e-07 -0.1437192 0.05694073 0.01160233
##   total_effect pval_total
## 1 0.04731923 0.2828259
```

`est_beta` is the estimated group-specific exposure-outcome effect for the reference group, and `se_beta` and `pval_beta` are its standard error and p-value, respectively. `est_beta_int` is the estimated exposure-group interaction effect, and `se_beta_int` and `pval_beta_int` are its standard error and p-value, respectively. `total_effect` is the estimated group-specific exposure-outcome effect for the comparison group and `pval_total` is its respective p-value.

## Example 2: Three-Sample Data

In this example, we perform an analysis on simulated three-sample data. The configuration is similar to the two-sample analysis, except the `data_list_3sample` argument is used instead of `data_list_2sample` and `model_type` is set to “3sample”.

```
result_3sample <- int2MR(data_list_3sample = example_3sample_data,
  model_type = "3sample",
  prior_inv_gamma_shape = 0.1,
  prior_inv_gamma_scale = 0.1,
  chains = 2, iter = 10000, warmup = 2500,
  adapt_delta = 0.95)
```

```
# Display the results for the three-sample analysis
result_3sample$result_3sample
```

```
##   est_beta   se_beta   pval_beta est_beta_int se_beta_int pval_beta_int
## 1 0.1981924 0.03203621 6.151069e-10 -0.1399829 0.05635217 0.01298891
##   total_effect pval_total
## 1 0.05820944 0.1102825
```