ISMB_Tutorial_Example

Tayyaba Alvi

2025-06-05

Overview

This hands-on session is a draft guide through the steps to perform Mendelian Randomization (MR) using R and publicly available GWAS summary statistics. The complete coding material will be available by end of June and will include detailed analysis and pre-processing steps. The final repo will contain a gitpod link which will have the R enviroment and all packages installed.

Learning Objectives

By the end of this tutorial, participants will:

- Access and prepare GWAS summary statistics.
- Select appropriate genetic instruments.
- Harmonize datasets for MR analysis.
- Perform standard MR methods.
- Interpret MR results and test robustness.
- Understand key assumptions and limitations.
- Apply MR to real-world, health-relevant examples.

Environment Setup

Participants will use a pre-configured Gitpod workspace with RStudio. All necessary packages are pre-installed:

- `TwoSampleMR`
- `MRInstruments`
- `tidyverse`
- `ieugwasr`
- `gwasvcf`
- `gwasglue2`
- `genetics.binaRies`
- `ggplot2`

No installation needed.

Tutorial Structure

Step 1: Setup & Data Exploration

Goal: Familiarize with GWAS resources and identify exposures and outcomes.

```
library(TwoSampleMR)
library(ieugwasr)
```

```r
# Explore available exposures
gwas_exposures <- available_exposures()
head(gwas_exposures)

# Search for CRP-related GWAS
crp_traits <- subset(gwas_exposures, grepl("C-reactive", trait))
print(crp_traits)
```

Exercise:

- Explore and compare traits.
- Find CHD and IL6 traits.

Step 2: Instrument Selection & Harmonization

Goal: Extract and harmonize summary data.

```r
# Extract SNPs for CRP (example exposure)
exposure_data <- extract_instruments(outcomes = "ebi-a-GCST002216")

# Extract outcome data for CHD (example outcome)
outcome_data <- extract_outcome_data(snps = exposure_data$SNP, outcomes = "ebi-a-GCST003116")

# Harmonize datasets
harmonized_data <- harmonise_data(exposure_data, outcome_data)
head(harmonized_data)
```

Exercise:

- Examine harmonization steps.
- Explore allele alignment.

Step 3: Perform MR Analysis

Goal: Estimate causal effect using multiple MR methods.

```r
# Run MR
mr_results <- mr(harmonized_data)
print(mr_results)

# Basic scatter plot
mr_scatter_plot(mr_results, harmonized_data)
```

Exercise:

- Compare IVW vs MR-Egger.
- Interpret causal estimate direction and magnitude.

Step 4: Sensitivity & Pleiotropy Checks

Goal: Evaluate robustness of MR results.

```r
# Heterogeneity
heterogeneity <- mr_heterogeneity(harmonized_data)
print(heterogeneity)

# Pleiotropy (Egger intercept)
pleiotropy <- mr_pleiotropy_test(harmonized_data)
print(pleiotropy)

# Leave-one-out analysis
loo <- mr_leaveoneout(harmonized_data)
mr_leaveoneout_plot(loo)
```

Exercise:

- Interpret heterogeneity p-values.
- Identify influential SNPs from LOO plot.

Step 5: Case Studies — CRP vs IL6

```r
# CRP: GCST002216, CHD: GCST003116
exposure_data_crp <- extract_instruments("ebi-a-GCST002216")
outcome_data_chd <- extract_outcome_data(exposure_data_crp$SNP, "ebi-a-GCST003116")
harmonized_crp <- harmonise_data(exposure_data_crp, outcome_data_chd)
results_crp <- mr(harmonized_crp)
mr_scatter_plot(results_crp, harmonized_crp)
```

Example 1: CRP → CHD

> **Conclusion**: Observational association between CRP and CHD appears confounded — MR shows no causal effect.

```r
# IL6: GCST90029070, CHD: GCST003116
exposure_data_il6 <- extract_instruments("ieu-a-300")  # Replace with correct ID if different
outcome_data_chd_il6 <- extract_outcome_data(exposure_data_il6$SNP, "ebi-a-GCST003116")
harmonized_il6 <- harmonise_data(exposure_data_il6, outcome_data_chd_il6)
results_il6 <- mr(harmonized_il6)
mr_scatter_plot(results_il6, harmonized_il6)
```

Example 2: IL6 → CHD

> **Conclusion**: IL6 shows a causal effect on CHD, supporting inflammatory pathway involvement.

Exercise:

- Compare MR results from CRP and IL6.
- Discuss biological plausibility.

Summary & Discussion

- MR requires strong assumptions: relevance, independence, and exclusion.
- Sensitivity analyses are essential to interpret causality.
- Contrasting examples (CRP vs IL6) highlight how MR can distinguish correlation from causation.