Joint selection of exposures and horizontal pleiotropy in multivariable Mendelian randomization with application to causal gene identification

Yihe Yang, Noah Lorincz-Comi, Xiaofeng Zhu

Department of Population & Quantitative Health Sciences, Case Western Reserve University

Background: Jointly analyzing genome-wide association studies (GWAS) data and expression quantitative trait loci (eQTL) data will potentially identify causal genes of diseases. However, current univariable methods such as colocalization and univariable Mendelian randomization (UVMR) may not be robust due to high correlations among gene expressions, horizontal pleiotropy, and biases because of weak instruments in MR analysis.

Methods: We propose a novel multivariable method named MR Joint Outliers-aNd-Exposures Selection (Mr.Jones), which performs multivariable MR (MVMR) analysis in a genome region using multiple gene expressions as exposures. Mr.Jones applies the unbiased estimating function to mitigate weak instrument bias, employs variable selection penalties to simultaneously select causal genes and identify horizontal pleiotropy, and decorrelates the instrument variables using their linkage disequilibrium (LD) matrix.

Results: In simulations, Mr.Jones resulted in unbiased causal effects estimates in the presence of horizontal pleiotropy and many weak instrument variables compared with current MR methods. We applied Mr.Jones to search causal genes for coronary artery disease (CAD), type 2 diabetes (TD2), and osteoarthritis (OA) in European populations, identifying *TCF7L2* as a protective gene for diabetes, *MTAP* as a risk gene for CAD, and *SLC25A1*3 as a risk gene for OA.

Conclusion: As more GWAS and eQTL data become publicly available, Mr.Jones can serve as a valuable tool in studying causal relationships between gene expressions and diseases, therefore enabling the further understanding of disease mechanisms, and facilitating precision medicine and drug development.