MR-LDP: a two-sample Mendelian randomization for GWAS summary statistics accounting linkage disequilibrium and horizontal pleiotropy

Jin Liu

Duke-NUS Medical School
E-mail: jin.liu@duke-nus.edu.sg

Abstract: The proliferation of genome-wide association studies (GWAS) has prompted the use of two-sample Mendelian randomization (MR) with genetic variants as instrumental variables (IV) for drawing reliable causal relationships between health risk factors and disease outcomes. However, the unique features of GWAS demand MR methods account for both linkage disequilibrium (LD) and ubiquitously existing horizontal pleiotropy among complex traits, which is a phenomenon that a variant affects the outcome other than exclusively through the exposure. Therefore, statistical methods that fail to consider LD and horizontal pleiotropy can lead to biased estimates and false-positive causal relationships. To overcome these limitations, we propose a probabilistic model for MR analysis to identify casual effect between risk factors and disease outcomes by using GWAS summary statistics in the presence of LD, as well as properly accounts for horizontal Pleiotropy among genetic variants (MR-LDP). MR-LDP utilizes a computationally efficient parameter-expanded variational Bayes expectation-maximization (PX-VBEM) algorithm, calibrating the evidence lower bound (ELBO) for a likelihood ratio test. We further conducted comprehensive simulation studies to demonstrate the advantages of MR-LDP over existing methods in terms of both type-I error control and point estimates. Moreover, we used two real exposure-outcome pairs (CAD-CAD and BMI-BMI; CAD for coronary artery disease and BMI for body mass index) to validate results from MR-LDP in comparison with alternative methods, particularly showing that our method is more efficient using all instrumental variants in LD. By further applying MR-LDP to lipid traits and BMI as risk factors on complex diseases, we identified multiple pairs of significant causal relationships, including protective effect of high-density lipoprotein cholesterol (HDL-C) on peripheral vascular disease (PVD), and positive causal effect of body mass index (BMI) on haemorrhoids.