Projected and constrained Mendelian randomization as a robust instrumental variable method for causal inference in high dimensional settings

Celia Greenwood*
celia.greenwood@mcgill.ca

In Mendelian randomization, genetic variants (SNPs) are used to construct an instrumental variable to estimate the causal effect of a phenotype (or exposure) on a disease. However, the existence of pleiotropy violates the assumption that the instrument (SNPs) and the response (disease) are independent conditional on the phenotype of interest. As a result, the ordinary two-stage least squares estimator, using all desired SNPs as instrumental variables, leads to biased estimation of the required causal effect.

We propose novel penalized and constrained methods to perform adjusted causal effect estimation as well as SNP selection, by finding a penalized projection orthogonal to a set of possibly pleiotropic phenotypes that are not of primary interest. Assuming that there are sufficient potential genetic instruments, constrained quadratic optimization with a smoothed-L0 norm can correct the bias induced by pleiotropy and lead to sparse models and stable instrumental variables.

In simulations, we compared our approach to a naive method (using all SNPs), the limited information maximum likelihood method, canonical correlation analysis (CCA/sparseCCA), constrained stepwise selection methods, and the inverse probability weight adjusted regression method. Results show our approach leads to causal effect estimators with the smallest bias and variance among those compared.

In conclusion, our approach finds a robust sparse model, enforces automatic feature selection, and leads to better estimate of causal effects even when pleiotropy is present.

*Department of Clinical Epidemiology, Lady Davis Institute for Medical Research, 3755 Côte Ste-Catherine, Montréal, QC H3T 1E2, CANADA