Augmented methods in PheWAS studies for pleiotropic effects using biobank data

Yong Chen

University of Pennsylvania
E-mail: ychen123@upenn.edu

Abstract: Pleiotropy, where 1 genetic locus affects multiple phenotypes, can offer significant insights in understanding the complex genotype–phenotype relationship. Although individual genotype–phenotype association shave been thoroughly explored, seemingly unrelated phenotypes can be connected genetically through common pleiotropic loci or genes. However, current analyses of pleiotropy have been challenged by both methodologic limitations and a lack of available suitable data sources.

In this talk, we present a regression framework, reduced rank regression, to simultaneously analyze multiple phenotypes and genotypes to detect pleiotropic effects. We used a large-scale biobank linked electronic health record data from the Penn Medicine BioBank to select 5 cardiovascular diseases (hypertension, cardiac dysrhythmias, ischemic heart disease, congestive heart failure, and heart valve disorders) and 5 mental disorders (mood disorders; anxiety, phobic and dissociative disorders; alcohol related disorders; neurological disorders; and delirium dementia) to validate our framework.

Compared with existing methods, reduced rank regression showed a higher power to distinguish known associated single-nucleotide polymorphisms from random single-nucleotide polymorphisms. In addition, genome-wide gene-based investigation of pleiotropy showed that reduced rank regression was able to identify candidate genetic variants with novel pleiotropic effects compared to existing methods.