MASAMB 2023

CoPheScan: A Bayesian PheWAS approach

Ichcha Manipur[1,2], Guillermo Reales[1,2], Jae Hoon Sul[3], Simonne Longerich[3], Adrian Cortes[4], Chris Wallace[1,2,5]

[1]Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, University of Cambridge, Cambridge CB2 0AW, UK

[2]Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge CB2 2QQ, UK

[3]Merck & Co., Inc., Rahway, NJ, USA

[4]GlaxoSmithKline, Stevenage SG1 2NY, UK

[5]MRC Biostatistics Unit, University of Cambridge, Cambridge, United Kingdom

Abstract

Phenome-wide association studies (PheWAS) are a promising tool for drug development as they enable detection of pleiotropy (additional causal effects on human phenotypes) by a focused study of genetic variants that have known causal effects on diseases or have a functional impact on molecular mechanisms. However, traditional PheWAS methods are unable to identify true pleiotropy due to their inability to account for confounding due to correlation between variants (linkage disequilibrium, LD).

We developed CoPheScan (Coloc adapted Phenome-wide Scan), a Bayesian approach to PheWAS, that enables the detection of causal associations while simultaneously addressing LD confounding. This is enabled by analysing the genomic region around the index variant rather than focussing on single variant associations as in conventional approaches. We also incorporated a hierarchical model to infer priors from data with the ability to include covariate information such as average shared genetic influence (genetic correlation) between index and test traits.

We show with results from simulated data that our method provides substantially better control of false positive rates compared to traditional PheWAS approaches. We performed PheWAS using CoPheScan on genetic variants known to truncate proteins, alter protein abundance or change disease risk in 2275 UK Biobank phenotypes. We identified complex associations of known pleiotropic genes such as *APOE* and potential novel associations such as rs214830, a genetic variant known to alter the abundance of *TGM3*, with skin cancer. We also show the effect of the level of functional enrichment of the tested variant sets on priors inferred using the hierarchical model.