Causal multi-trait analysis of Genome-Wide Association Studies data

Azza Ahmed[1,2], Willem Kruijer[3], Ernst Wit[4] Marco Grzegorczyk[1]

[1]Bernoulli Institute, University of Groningen, Groningen, the Netherlands
[2]Center for Bioinformatics and Systems Biology and Department of Electrical and Electronic Engineering, University of Khartoum, Khartoum, Sudan
[3] Biometris, Wageningen University and Research, Wageningen, the Netherlands
[4]Institute of Computational Science (ICS) and Social Network Analysis Research Center (SoNAR-C), Universita della Svizzera italiana, Lugano, Switzerland

E-mail for correspondence: xxx@xxx.xx

Abstract: Univariate linear mixed models are commonly used in the analysis of genome-wide association studies (GWAS). However, with the rise of deeply phenotyped data in population-scale biobanks, there are now better opportunities for understanding relations between traits (in light of their genetic and non-genetic causes) by employing multivariate approaches. Causal inference (from observational data) of directed biological networks between traits and their causes lead to a better understanding of disease etiology, identification of risk factors, and ultimately, more efficient diagnosis and treatment options.
This talk gives an extension to a causal discovery framework, pcgen (Kruijer et al. 2020), that employs the PC algorithm for building a network between traits and genetic causes. An overview of pcgen will be provided, along with extensions: 1) to make it applicable to human studies by accounting for genetic relatedness, 2) based on individual-level genotype data from a homogeneous cohort, on whom multiple phenotypes measurements are available, 3) at a reduced computational cost. A case study applying this approach to a human GWAS dataset will also be presented, where direct genetic and non-genetic effects between traits are delineated.

Key words: GWAS; Causality; pc algorithm, multivariate analysis

Kruijer W, Behrouzi, P Bustos-Korts D, et al (2020). Reconstruction of networks with direct and indirect genetic effects Genetics: 214(4), 781–807