

Bayesian High-Dimensional Biological Pathway-Guided Mediation Analysis With Application To Metabolomics

Yuzi Zhang

Abstract:

With advances in high-resolution mass spectrometry technologies, metabolomics data are increasingly used to investigate biological mechanisms underlying associations between exposures and health outcomes in clinical and epidemiological studies. Mediation analysis is a powerful framework for investigating a hypothesized causal chain and when applied to metabolomics data, a large number of correlated metabolites belonging to interconnected metabolic pathways need to be considered as mediators. The identification of metabolic pathways as mediators is of great interest to gain biological insights into exposure-outcome relationships. However, existing approaches typically focus on identifying individual metabolites, which may limit the understanding of biological processes underlying exposure-outcome relationships. We propose a Bayesian biological pathway-guided mediation analysis that aims to identify metabolic pathways directly as active mediators and characterize metabolic pathway-specific indirect effects. This is accomplished by incorporating existing biological knowledge of metabolic pathways to account for correlations among mediators, along with variable selection techniques. We apply the proposed method to examine the role of metabolism in mediating the effect of prenatal exposure to per- and polyfluoroalkyl substances (PFAS) on gestational age at birth. Our analyses confirm metabolic pathways previously identified and provide additional uncertainty quantification.