Integrating Multiple Correlated Phenotypes for Genetic Association Analysis by Maximizing Heritability

Jin Zhou
jzhou@email.arizona.edu
Department of Biostatistics, University of Arizona, USA

Many correlated disease variables are analyzed in genetic studies of complex diseases in the hope of increasing power to detect causal genetic variants. A common statistical approach involves assessing the relationship between each phenotype and each single nucleotide polymorphism (SNP) individually and using a Bonferroni correction for the effective number of tests conducted. Alternatively, one can apply a multivariate regression or a dimension reduction technique, such as principal components analysis, and test for the association with the principal components of the phenotypes rather than the individual phenotypes. Other previous approaches have developed methods for combining phenotypes to maximize heritability at individual SNPs. These approaches are not practical for population samples with genome-wide scans. In this paper, we construct a maximally heritable phenotype, which is a linear combination of the various phenotypes by taking advantage of the estimated heritability and co-heritability. Our approach estimates heritability globally and is therefore applicable to genome-wide scans. Theoretically, and through simulations, we compare our approach with commonly used methods and assess both the heritability and the power of the overall phenotype. Our approach always increases heritability over the individual phenotypes being combined and has higher power for SNPs explaining a percentage of heritability, relative to using a single phenotype. Moreover we provide suggestions for how to choose the phenotypes for combination. Applications of our approach to a COPD genome-wide association study show the practical relevance.