Feasibility of using COSA as a genome-wide SNP screen

Gitta Lubke[1*] **, Jacqueline Meulman**[2]

1. University of Notre Dame, USA
2. Leiden University, The Netherlands

* Contact author: glubke@nd.edu

Keywords: COSA, genome-wide association, SNP

Genome-wide association studies (GWAS) of psychiatric disorders have not yet resulted in a significant breakthrough. The common approach in GWAS is to test single nucleotide polymorphisms (SNPs) univariately, thus necessitating substantial corrections for multiple testing. The complexity of the phenotype is reduced to a binary indicator (i.e., case vs. control). The Clustering Objects on Subsets of Attributes (COSA) algorithm is designed to detect clusters of objects (here: subjects) that are similar to each other, and to simultaneously select cluster-specific subsets of attributes (here: SNPs and phenotype items) that are relevant for the clusters. COSA permits a joint analysis of SNPs and phenotype items, and therefore combines the advantages of a multivariate analysis with the flexibility of targeting phenotype symptom patterns and clusters of subjects within the cases. COSA has not been used for the analysis of genome-wide SNP data. We present results of a simulation study that investigates the feasibility of using COSA as a genome-wide SNP screen.

References

Friedman, JF, Meulman, JJ (2004). Clustering objects on subsets of variables (with discussion). Journal of the Royal Statistical Society, Series B, 66, 815–849.