

ON USING PRIOR INFORMATION TO IMPROVE POWER OF GENOME-WIDE STUDIES: WEIGHTED P-VALUE VERSUS STRATIFIED FALSE DISCOVERY CONTROL

Yun Joo Yoo, Lei Sun, Shelley B. Bull

*Department of Public Health Sciences, University of Toronto, Toronto, Canada
Programs in Genetics and Genomic Biology, Hospital for Sick Children, Toronto, Canada
Department of Statistics, University of Toronto, Toronto, Canada
Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Canada*

Multiple hypothesis testing is one of major issues that can affect the success of genome scale studies. Exhaustive searches of hundreds of thousands of markers to detect true signals will require more stringent threshold resulting in significantly reduced power. The power of genome-wide studies, however, can be improved by utilizing prior information. Two non-Bayesian methods are available to incorporate prior information into hypothesis testing procedures in the framework of false discover rate. One is weighted p-value method and the other is stratification method. We investigate these two methods theoretically and evaluate their performances in power through simulation.

Theoretically, we show that the stratification method can be formulated as a weighted p-value method, where each stratum is assigned with a fixed weight across all markers within the stratum. However, the weighting factors for stratification method is internally defined sparing the effort of determining specific weights. To investigate the power, we simulated both informative and uninformative genome-wide linkage signals with corresponding SNP association data with 10 or 100 true SNPs. The weighting or the stratification schemes were determined using the information from the linkage signals. The simulation results show that when the prior is not informative, unlike the weighted p-value method, the stratification method does not lose power. When the prior is informative, both methods improve power with a marginally better performance by the weighted p-value method.

Roeder K, Bacanu S., Wasserman L, Devlin B (2006) Using linkage genome scans to improve power of association in genome scans. *Am J Hum Genet* 78:243-252.

Sun L, Craiu RV, Paterson AD, Bull SB (2006) Stratified false discovery control for large scale hypothesis testing with application to genome-wide association studies. *Genet Epidemiol* 28:352-367.