

Optimal two-stage genotyping in association studies for complex diseases

Jaya Satagopan, Ph.D.
Memorial Sloan-Kettering Cancer Center

Abstract:

Gene-disease association studies based on case-control designs are often used to identify markers conferring disease risk. When examining a large number of markers, as in whole-genome screening or large candidate marker studies, a one-stage approach that evaluates all the markers on all samples may be inefficient in resource utilization. This talk addresses a cost-effective two-stage approach to investigate gene-disease associations, where all the markers are first evaluated on a fraction of the available subjects at Stage 1 and the most promising markers are then evaluated on the remaining subjects at Stage 2. The goal is to determine the fraction of individuals to be utilized in Stage 1 and the fraction of the most promising markers to be evaluated in Stage 2. The properties of the two-stage design are evaluated under various parametric configurations. The results show that, compared to a one-stage design, a two-stage approach typically halves the cost of the study. As a general guideline, the simulations indicate that evaluating all the markers on 50% of the individuals at Stage 1 and evaluating the most promising 10% of the markers on the remaining individuals at Stage 2 provides near-optimal power while resulting in a 45% decrease in the total number of marker evaluations.