Case-control studies have been extremely valuable in evaluating associations between candidate genes and complex diseases. Traditional case-control studies use unrelated subjects and compare allele or genotype frequencies of the cases and the controls at genetic markers. When affected related individuals are used in association studies, the power to detect an association is increased since affecteds with affected relatives have a higher expected frequency of the alleles that increase susceptibility for a genetic trait than do affected individuals that do not have affected relatives. When related individuals are used in a study, the correlations among the relatives must be taken into account to ensure validity of the test, and consideration of these correlations can also improve power. This thesis focuses on the problem of testing for association between a binary trait and a genetic marker when cases and/or controls are related. The emphasis is on methods that are computationally feasible even for hundreds of thousands of markers and/or extremely complex pedigrees.

Case-control association testing is essentially a comparison of the allele frequency distribution between cases and controls. We first consider the related problem of allele frequency estimation, which is important in its own right because many genetic mapping and population genetic analyses require allele frequency estimates. We propose a new estimator, which is an extension of the best linear unbiased estimator of McPeek et al. (2004), and which uses a random covariance matrix that depends on the observed genetic data at loci that are linked to the locus of interest to increase efficiency. Our estimator is unbiased if the linked loci are in linkage equilibrium with the locus of interest and it is not as computationally intensive as the MLE.

To test for allelic association with a binary trait, we propose a new test, the $M_{QLS}$ test, that is an extension of the quasi-likelihood score test of Bourgaín et al. (2003) and which takes advantage of the fact that affected individuals that have affected relatives are more likely to have the predisposing variants than individuals that do not have affected relatives. One of the motivations for using the $M_{QLS}$ test is that for any arbitrary set of outbred individuals, the $M_{QLS}$ statistic has maximal noncentrality parameter in a general class of linear statistics, for all 2 allele disease models, as the effect size tends to zero. We perform simulations to compare the type I error and the power of the $M_{QLS}$ and competing methods. We apply the methods to analyze data on asthma-related phenotypes in a complex Hutterite pedigree (in collaboration with Dr. Carole Ober) and an alcoholism related phenotype in a sample of moderate size outbred Caucasian pedigrees from GAW 14 COGA data.

Information about building access for persons with disabilities may be obtained in advance by calling the department office at (773) 702-8333.