UCLA

BIOSTATISTICS SEMINAR

SPRING 2012

From Local Ancestry to Rare Variant Analysis: Some Current Issues in Genetic Association Studies

David V. Conti, PhD

Associate Professor

Department of Preventative Medicine, Division of Biostatistics

University of Southern California

Candidate for the Faculty position in Biostatistics-Epidemiology Departments

Monday, April 23, 2012

3:30pm - 4:30pm, CHS 33-105A

Refreshments served at 3:00 PM in room 51-254 CHS

ABSTRACT: Current genetic association studies face many challenges as they move into the "post-GWAS" era. These include the impact of local ancestry, the design of sequencing studies, and the analysis of rare variants. In this talk, I first discuss the mechanisms in which admixture may lead to confounding by population substructure and heterogeneity due to differential patterns of linkage disequilibrium (LD). Understanding these mechanisms has important implications for the performance of association tests and I demonstrate how the use of local ancestry can help to highlight novel findings in the University of Southern California's Children's Health Study investigating asthma risk. Next, I discuss issues involving the design and analysis of association studies using next-generation sequencing. While the ability to identify individual-level data is lost (without bar-coding), sequencing pooled samples can have many design advantages for both discovery and association testing. For pooled data, I present a hierarchical Bayesian modeling approach that estimates the association of each variant using pools of cases and controls while accounting for the variation in read depth across pools and sequencing error. I discuss how the optimal design and performance is influenced by the number of pools, the number of individuals within each pool, and the average coverage per pool. This approach is then contrasted with a novel approach for rare variant analysis for individual-level data. Here, a Bayesian model uncertainty approach is used to average over the inclusion and direction of effect for each variant in a risk index. The approach allows for inference at both the group and variant-specific levels and shows increased power over alternative rare variant analysis methods. Future design and analytic approaches in "post-GWAS" genetic association studies must be cognizant of the interplay between many underlying mechanisms. I end the talk by discussing methodological extensions for accounting for local ancestry in rare variant analysis and the incorporation of prior biological information.