



JOHNS HOPKINS  
BLOOMBERG  
SCHOOL of PUBLIC HEALTH

*Department of Biostatistics*

# BIOSTATISTICS SEMINAR

## A Non-parametric Approach to Multi-SNP GWAS

**Dr. Knut M. Wittkowski**  
**The Rockefeller University**

### Abstract

When U-statistics were originally developed in the 1940s, the theory included applications to multivariate data, in general, yet the absence of computers restricted their use to methods for uni- or bi-variate data, such as sign, Mann-Whitney, and Gehan test. When sufficiently powerful computers became more widely available, another obstacle to a wider use emerged. Being based on counting pairwise orderings, information content (the proportion of unambiguous orderings) decreased dramatically with an increase in the number of variables included.

Here we propose strategies how prior information about the variables, ie, their relevance order or factor structure can be utilized to increase information content without requiring that assumptions about additivity and linearity be made. We will discuss examples where such prior information comprises neighborhood on a chromosome, degree of severity or relatedness (side effects), or etiology (phenotypes).

Having a more powerful multivariate approach available and also the grid technology to enable it for identifying epistasis in common diseases and their impact on complex phenotypes has implications for the HapMap project and has potential to advance comparative effectiveness research and personalized diagnostics or treatments.

**The Johns Hopkins Bloomberg School of Public Health**  
**Department of Biostatistics, Wednesday, December 8, 2010**  
**Room W2030 School of Public Health, 4:00-5:00pm**