

**Jill Barnholtz-Sloan, PhD, Assistant Professor, H. Lee Moffitt Cancer Center
and Research Institute, Tampa, Florida**

**Using Race/Ethnicity and Individual Ancestry as Tools in Genetic
Association Studies**

Abstract:

Human phenotypic variation is more common between individuals residing on the same continent than between individuals from different continents. This variation is related to how ancestral populations have mixed with each other over time, i.e. population stratification. Population stratification has the potential to affect the results of genetic marker studies. Estimating individual ancestry provides a continuous measure to assess population structure in case-control studies of complex disease, instead of using self-reported racial groups. Within any self-reported racial group in the United States there is a significant amount of heterogeneity by ancestry. In order to estimate ancestry it is necessary to first genotype a panel of ancestry informative markers and then the genotypes from these markers can then be used to estimate ancestry using various mathematical algorithms. Examples using this technique with various phenotypes will be presented.

In addition, we present our work using individual ancestry as a surrogate for self-reported race to investigate an association between having the *GSTM1* null risk genotype and early-onset lung cancer risk from a case-control study of African Americans and Caucasian, non-Hispanics in Detroit, Michigan (Barnholtz-Sloan et al, CEBP, 2005). Individual ancestry proportions were estimated for "European" and "West African" groups using published allele frequencies. The majority of Caucasian, non-Hispanics had greater than 50% "European" ancestry while the majority of African Americans had lower than 20% "European" ancestry, although significant overlap by self-reported race and ancestry also existed. We found that the frequency of the *GSTM1* null genotype varies by individual "European" ancestry and case-control status within self-reported race (particularly for African Americans). Genetic risk models showed that adjusting for individual "European" ancestry provided a better fit to the data compared to the model with no group adjustment or adjustment for self-reported race. This study suggests that significant population substructure differences exist that self-reported race alone does not capture and that individual ancestry may be confounded with disease status and/or a candidate gene risk genotype.