# Association Analysis under Population Stratification: A Two-Stage Procedure Utilizing Population- and Family-Based Analyses

Hui-Wen Lin
Biostatistics Research and Consulting Center,
Taipei Medicine University

## Abstract

The association analysis based on a population-based case-control study is convenient and powerful, but may be biased under population stratification (PS), namely the study population consists of strata heterogeneous in disease rates and allele frequencies. On the other hand, a family-based (e.g. case-parents) study is robust against the PS bias, but may be less convenient to implement. We propose an association analysis that preserves the full robustness property of the family-based analysis while allowing for borrowing information from a population-based analysis. We use two-stage procedure for our method. In the first stage, one selects a population-based case-control sample and performs a traditional case-control association analysis. In the second stage, one randomly selects a subset of the first-stage cases and recruits their family controls (e.g. parents), and performs a family-based association analysis. An overall two-stage analysis is then performed to utilize information from the two stages. The proposed two-stage analysis achieves higher power than the second-stage family-based analysis by utilizing information in the first-stage population study, while is always valid even under PS. The proposal can further accommodate parental missingness when the case-parents study is used as the second-stage family study. The computation- and cost-effectiveness render the new proposal very promising in genome-wide association studies. This is a joint work with Dr. Yi-Hau Chen.