Evaluating the power of single variant association tests for low frequency variants in joint and meta-analysis

Clement Ma, Tom Blackwell, Michael Boehnke and Laura J. Scott

Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI

In case-control association testing of common genetic variants (minor allele frequency [MAF] ≥ 5%), combining logistic regression results across multiple studies using fixed effects meta-analysis has comparable power to joint analysis of the combined individual-level data (Lin & Zeng, 2010). However for less common variants (MAF < 5%), the relative power of joint and meta-analysis, and the most powerful and best-calibrated single variant test are unknown. In this study, we use analytic calculation and computer simulation to compare the type I error rate and power of the (1) Wald, (2) score, (3) likelihood ratio [LR], and (4) Firth bias-reduced logistic regression case-control association tests in both (a) joint and (b) meta-analysis for variants with low MAF or more precisely, low minor allele count [MAC].

For joint analysis of less common variants, the Firth test consistently has the best-calibrated type I error rate in both balanced (equal numbers of cases and controls) and unbalanced studies. The score test is conservative for balanced studies, but becomes increasingly anti-conservative with increasing case-control imbalance. In general, meta-analysis does not have comparable power than joint analysis as MAC decreases. Score test-based meta-analysis is best-calibrated in balanced studies, but is severely anti-conservative in unbalanced studies. Meta-analysis of Firth test results is conservative in balanced studies, and is anti-conservative in unbalanced studies. Power mostly correlates with the relative conservativeness of each test. If the combined individual-level data are available, for association testing of less common variants in both balanced and unbalanced studies we recommend joint analysis using the Firth test. If individual-level data are unavailable, in balanced studies we recommend meta-analysis of score test results.