

CONTROL ID: 3469501

PRESENTER: Wonil Chung

PRESENTER (INSTITUTION ONLY): Soongsil University, Harvard T.H. Chan School of Public Health

TITLE: Efficient Penalized Regression Approaches Improve Polygenic Prediction in Biobank data

ABSTRACT BODY:

Abstract Body: We introduce CTPR (Cross-Trait Penalized Regression), a powerful and practical approach for multi-trait polygenic risk prediction in Biobank-scale cohorts. Specifically, we propose a novel cross-trait penalty function with Lasso and MCP to incorporate the shared genetic effects across multiple traits for large-sample GWAS data. Our approach extracts information from the secondary traits that is beneficial for predicting the primary trait based on individual-level genotypes and/or summary statistics. Our novel implementation of a parallel computing algorithm makes it feasible to apply our method to biobank-scale GWAS data. We illustrate our method using large-scale GWAS data (~1M SNPs) from the UK Biobank (N=456,837). We show that our multi-trait method outperforms the recently proposed multi-trait analysis of GWAS (MTAG) for predictive performance. The prediction accuracy for height by the aid of BMI improves from $R^2=35.8\%$ (MTAG) to 42.5% (MCP+CTPR) or 42.8% (Lasso+CTPR) with UK Biobank data. Next, we extend our CTPR method to multi-ethnic GWAS data by modelling population-specific LD. Assuming similar LD structure around a SNP between two distinct populations leads to similar coefficients of the SNP, the SNP specific adjacency coefficient in the cross-ethnic penalty is defined as functional form of LD score. This function can improve the prediction accuracy for complex traits of non-European ancestry (primary ethnicity) using European ancestry (secondary ethnicity) with large samples. Furthermore, it can take advantage of the secondary ethnicities based on summary statistics and thus have the potential to utilize information from most published GWAS summary statistics. The predictive performance of CTPR (Cross-ethnic Penalized Regression) will be compared with the existing multi-ethnic prediction methods.

AUTHORS/INSTITUTIONS: W. Chung, Statistics, Soongsil University, Seoul, KOREA (THE REPUBLIC OF)|W. Chung, L. Liang, Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, UNITED STATES|