Use of GWAS by subtraction in lieu of covariates: the expected and unexpected.

Katrina L. Grasby

1 QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

KEYWORDS: GWAS, GWAS-by-subtraction, method comparison, ENIGMA

ABSTRACT:

GWAS summary statistics are extremely useful for a myriad of follow up analyses, which are made possible through various open-access programs, including genomicSEM. With large collaborations of researchers pooling their data and making GWAS results freely available, there is an ever increasing amount of data available to improve our understanding of genotype and phenotype. In a recent GWAS of the cortex by the ENIGMA consortium, in excess of 11 TB of data were uploaded to a central server for analysis. However, approximately half of the data was essentially a duplicate model that included a global control measure. I explored if GWAS-by-subtraction would provide equivalent results, thus requiring half the GWAS and half the data storage for future work that implement similar protocols. From the cortical results, GWAS-by-subtraction was run “subtracting” total surface area from the genetic effects on the surface area of the precentral sulcus. Results were compared with the original GWAS that included total surface area as a covariate in the regression. The GWAS-by-subtraction results genetically correlated 1.0 with the original summary statistics. However, only a third of the original loci were significant. While most loci were slightly reduced in significance, one very strong signal was reduced by a large amount. Reasons for this surprising result will be discussed.


GRANT SUPPORT: Australian NHMRC APP1173025