Supplementary Note for: Genetic correlation, pleiotropy, and causal associations between substance use and psychiatric disorder

Seon-Kyeong Jang, M.A.1, Gretchen Saunders, Ph.D. 1, MengZhen Liu, Ph.D. 1, 23andMe Research Team, Yu Jiang, M.S.2,3, Dajiang J. Liu, Ph.D.2,3, Scott Vrieze, Ph.D.1

1Department of Psychology, University of Minnesota, Minneapolis, MN

2Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA

3Institute of Personalized Medicine, Penn State College of Medicine, Hershey, PA

Contents

1. Quality control procedure for GWAS studies

2. Bidirectional Mendelian Randomization and Latent Causal Variable

3. Two-sample Mendelian Randomization

**1. Quality control procedure for GWAS studies**

Variants with the following characteristics were removed: 1) MAF < 0.01 to ensure that only well-imputed common variants are included, 2) when per-variant sample sizes were reported, we removed those variants with sample size of 50% of the mean sample size across all variants for that study, 3) when imputation scores were reported, we removed those with quality scores < 0.8, 4) multi-allelic sites, 5) variants in the MHC region due to complex linkage disequilibrium (LD) patterns (chr6:28477797-33448354), and 6) indels due to the possibility of inconsistent allele coding across studies. Reference and alternate alleles were checked against the Haplotype Reference Consortium (HRC) reference panel (McCarthy *et al.*, 2016) and any missing allele frequencies were copied from the HRC (e.g., to inform the MAF filter above). For LD Score Regression (LDSC) and genomic SEM (Bulik-Sullivan *et al.*, 2015), we further filtered to include only variants polymorphic in the European subset of HapMap3. To obtain genome-wide sentinel variants (i.e., the variant with lowest p-value within a region), we used Priority Pruner v0.1.4 to LD clump variants with p-values less than 5×10-8 (r2 =0.1, maximum distance=50Kb, HRC genotype). These variants were also taken forward to the bivariate Mendelian randomization analysis.

**2. Bidirectional Mendelian Randomization** **and Latent Causal Variable**

For a given pair of phenotypes, two sets of correlations were calculated, first using GWS variants of the first phenotype and then repeating the same procedure using GWS variants for the second phenotype. Based on these correlations, we calculated Akaike Information Criteria (AIC) for four models: in model 1 and 2, phenotype A or B causes B or A; in model 3 there are no causal relationships; in model 4, two phenotypes are very closely related (e.g., two alternate measurement of single entity or one could be the major highly penetrant causal factor for the other). Please see Supplmentary Note 1.3 of Pickrell *et al.* 2016 for likelihood calculation for each model. If GWS variants of one phenotype do not exist in summary statistics of the other phenotype, they were replaced by another variant with highest LD within 500KB region in HRC European reference panel. Based on AIC, the relative fit was defined as equal to exp((*AICmod1,2*–AIC*mod3,4*)/2) where AIC*mod1,2*is the smallest AIC from the two causal models (Model 1 and 2) and AIC*mod3,4* isthe smallest AIC from the two non-causal models (Model 3 and 4). A relative fit <0.01 was interpreted as supporting the causality.

The Latent Causal Variable (LCV) method was performed using the LCV package in R 3.6.0. Previously computed European ancestry LD scores were used for all analyses (https://data.broadinstitute.org/alkesgroup/LDSCORE/). The number of jackknife blocks was set to 100, and both the LDSC and cross-trait intercepts were freely estimated (i.e., they were not fixed). Evidence for causal relationships between traits is based on an absolute value of the GCP > 0.6 combined with a p-value surviving Bonferroni correction.  **3. Two-sample Mendelian Randomization**

Two-sample MR was performed on 30 substance use × psychiatric disorder pairs (cross-category). Two-sample MR was not performed among substance use or psychiatric disorders within each category because substantial degree of sample overlap is suspected within category based on sample documentation of each study. Variants polymorphic in the European subset of HapMap3 and having *p*-values less than 5×10-8 were selected as instruments for exposure variables. These variants were further LD-clumped (r2<0.001). Among these variants, those which did not pass Steiger filtering were removed to prevent reverse causation (Hemani *et al.*, 2017). Mean F-statistics were calculated by average of $\hat{β}^{2}/ \hat{σ}^{2}$ for instrument variants (Lee, Goddard, Wray, & Visscher, 2012) to assess instrument strength (Pierce, Ahsan, & VanderWeele, 2011). Four widely used MR methods were applied on the exposure-outcome harmonized data, i.e., inverse variance weighted (IVW), MR Egger, weighted median, and weighted mode. To assess instrument heterogeneity which indicates possible presence of horizontal pleiotropy, Cochran’s Q in IVW, Rucker’s Q and MR Egger intercepts in MR Egger analysis were calculated. The full MR results including beta, SE, *p*-values as well as mean *F*- and heterogeneity statistics were reported in Supplementary Table 13.

**References**

Hemani, G., Tilling, K., & Smith, G. D. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genetics*, *13*(11), e1007081. doi: 10.1371/journal.pgen.1007081

Lee, S. H., Goddard, M. E., Wray, N. R., & Visscher, P. M. (2012). A better coefficient of determination for genetic profile analysis. *Genetic Epidemiology*, *36*(3), 214–224. doi: 10.1002/gepi.21614

Pickrell, J. K., Berisa, T., Liu, J. Z., Ségurel, L., Tung, J. Y., & Hinds, D. A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. *Nature Genetics*, *48*(7), 709–717.

Pierce, B. L., Ahsan, H., & VanderWeele, T. J. (2011). Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *International Journal of Epidemiology*, *40*(3), 740–752. doi: 10.1093/ije/dyq151