

Submitter Name: Silviu-Alin Bacanu  
Submitted email: silviu-alin.bacanu@vcuhealth.org  
PI Name: Christos Chatzinakos  
PI email: [Chris.Chatzinakos@vcuhealth.org](mailto:Chris.Chatzinakos@vcuhealth.org)

**Direct imputation of summary statistics using large reference panels increases the resolution of genetic signals coming from metanalysis of substance use disorder**

Christos Chatzinakos<sup>1</sup>; Vladimir I. Vladimirov<sup>1</sup>; Tim Bigdeli<sup>2</sup>; Bradley T. Webb<sup>1</sup>; Brien P. Riley<sup>1</sup>; Kenneth S. Kendler<sup>1</sup>; Silviu-Alin Bacanu<sup>1</sup>

<sup>1</sup>Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University;  
<sup>2</sup>SUNY Downstate Medical Center, University Hospital of Brooklyn, New York

To increase the signal resolution for large-scale meta-analyses of genome-wide association studies, genotypes at unmeasured Single Nucleotide Polymorphisms (SNPs) are imputed using multi-ethnic reference panels. Direct imputation of summary statistics was shown to be practically as accurate as genotype imputation method, while incurring orders of magnitude lower computational burden. However, to be precise, direct imputation methods need larger reference panels, whereas current methods use a small reference panel coming from the 1000 Genome project. A previous version of our software, DISTMIX, achieved such estimates by using in-cohort allele frequency (AF) estimates. However, due to privacy issues, genome wide association studies (GWAS) stopped providing AFs.

To accurately estimate the LD needed for analysis of mixed ethnicity cohorts, we propose a novel method, DISTMIX2, and, for a larger breadth of imputation, a much larger reference panel. The ethnic weights are now estimating using only Z-score summary statistics. Simulations show that the proposed method adequately controls the false positive rates. For a larger reference panel, we use the publicly and privately available data to obtain a 33,000 (33K) panel, which includes 11K Han Chinese.

Imputation of summary statistics from the latest substance use disorder meta-analyses uncover several new signals, some of which pinpoint potential treatment avenues.