

Name: Sandra Sanchez-Roige

Email: sanchezroige@ucsd.edu

Genome-Wide Association Analysis of an in Silico Phenotype Sourced from Clinical notes: The Addiction Behavior Checklist

Ellen Tsai¹, Parker Christy², Dagmawi Negesse², Lori Schirle², Angus Chatham², Eli Bradley², David Samuels³, Lea Davis^{3,4,5,6,7,8,9,10}, Alvin Jeffery², Sandra Sanchez-Roige^{6,11,12}

¹Department of Medicine, Icahn School of Medicine at Mount Sinai; ²School of Nursing, Vanderbilt University; ³Vanderbilt Genetics Institute, Vanderbilt University Medical Center;

⁴Department of Biomedical Informatics, Vanderbilt University Medical Center

⁵Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center;

⁶Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center;

⁷Department of Biomedical Informatics, Vanderbilt University Medical Center; ⁸Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center; ⁹Department of Genetic and Genomic Sciences, Icahn School of Medicine at Mount Sinai; ¹⁰Department of Psychiatry, Icahn School of Medicine at Mount Sinai; ¹¹Department of Psychiatry, University of California San Diego; ¹²Institute for Genomic Medicine, University of California San Diego

Background: Electronic health records (EHR) offer a rich source of longitudinal data, useful for increasing sample sizes and power in gene discovery. OUD detection in EHR typically relies on ICD codes, but these codes may lack sensitivity and positive predictive value due to underutilization. To address this, we expanded EHR searches to include clinical notes and derived scores from the Addiction Behavior Checklist (ABC), an established tool for assessing prescription opioid addiction.

Results: We generated ABC scores for 63,207 patients (European N = 52,930; African N = 10,277). SNP-heritability was estimated at $4.1 \pm 0.5\%$ (European) and $10.3 \pm 2.8\%$ (African). While no significant associations were found in the African cohort, a risk locus (rs4364183, $P = 2.11 \times 10^{-8}$) in SATB1-AS1 on chromosome 3 was identified in the European cohort. This locus has been linked to smoking, risk-taking behavior, and alcohol use. Polygenic scores derived from GWAS of opioid use and externalizing traits were significantly associated with ABC scores ($P > 7.8 \times 10^{-11}$). ABC PGS did not predict white blood cell count ($P = 3.3 \times 10^{-1}$), a negative control.

Conclusion: ABC scores derived from clinical notes offer a cost-effective strategy to advance genetic studies of OUD.