

Submitter Name: Andrew W Bergen
Submitted email: abergen@ori.org
PI Name (if different): James W Baurley and Andrew W Bergen
PI email (if different): baurley@biorealm.ai

Multiethnic Prediction of Nicotine Biomarkers and Associations with Nicotine Dependence

Andrew W. Bergen PhD^{1,2,*}, Christopher S. McMahan PhD³, Stephen McGee MS², Carolyn M. Ervin PhD², Hilary A. Tindle, MD, MPH^{4,5}, Loïc Le Marchand MD, PhD⁶, Sharon E. Murphy PhD⁷, Daniel O. Stram PhD⁸, Yesha M. Patel MS⁸, Sungshim L. Park PhD⁶, James W. Baurley PhD²

¹ Oregon Research Institute; ² BioRealm, LLC; ³ School of Mathematical and Statistical Sciences, Clemson University; ⁴ Department of Medicine, Vanderbilt University Medical Center; ⁵ Veterans Health Administration-Tennessee Valley Healthcare System Geriatric Research, Education and Clinical Center; ⁶ Cancer Epidemiology and University of Hawaii Cancer Center, University of Hawai'i; ⁷ Biochemistry, Molecular Biology, and Biophysics and Masonic Cancer Center, University of Minnesota; ⁸ Department of Preventive Medicine, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California;

Background: The nicotine metabolite ratio and nicotine equivalents are biomarkers of nicotine metabolism rate and intake. Genome-wide prediction of these biomarkers will extend biomarker studies to cohorts without measured metabolites, enable tobacco-related research, and may assist in smoking cessation therapy.

Methods: We screened genetic variants genome-wide for the urinary nicotine metabolite ratio (uNMR) and creatinine-standardized total nicotine equivalents (TNE), in 2,239 current cigarette smokers in five ethnic groups participating in the Multiethnic Cohort Study. We applied statistical learning algorithms to build prediction models on top-ranked genetic variants and non-genetic variables for each biomarker. We predicted these nicotine biomarkers using model ensembles in the training sample (internal validity). We used genome-wide data to predict nicotine biomarkers in 1,864 treatment-seeking smokers in two ethnic groups and assessed association with dependence measures (external validity).

Results: The genomic regions with the most selected and trained variants for measured biomarkers were chr19q13.2 (uNMR), and chr15q25.1 and chr10q25.3 (TNE). We observed correlations (measured and predicted) of 0.67 and 0.68 for the uNMR and 0.65 and 0.72 TNE in the training sample. In treatment-seeking smokers, predicted uNMR was significantly associated with CPD, and predicted TNE was associated with CPD, Time-To-First-Cigarette, and Fagerström total score.

Discussion: Nicotine metabolites, genome-wide data and statistical learning approaches develop novel robust predictive models for urinary nicotine biomarkers in a multiethnic sample. Predicted biomarker associations define genetically-influenced components of nicotine dependence. Secondary analysis of smoking cessation trial data will provide opportunities to assess translational relevance for application to tobacco use disorder treatment.

Financial Support This work was supported by National Institute on Alcohol Abuse and Alcoholism (R44 AA027675) and by the National Cancer Institute (R01 CA232516, U01 CA164973 and P01 CA138338).