Submitter Name: Stephen McGee
Submitted email: sjmcgee@biorealmresearch.com
PI Name (if different): James W Baurley PhD
PI email (if different): baurley@biorealm.ai

From the developers of Smokescreen: a booster panel for development of addiction polygenic risk scores

Stephen McGee MS[1], Andrew W. Bergen PhD[1,2], James W. Baurley PhD[1]

[1] BioRealm, LLC; [2] Oregon Research Institute

Background: Individuals suffering from addiction-related disorders have been immensely impacted by the COVID-19 pandemic. Given that both addiction and response to SARS-CoV-2 are partially driven by genetic factors, we propose a booster genotyping panel to help foster a better understanding of the genetics of addictions and COVID-19, develop polygenic risk scores (PRS), and help develop treatments for this vulnerable population.

Methods: We are refining our existing Smokescreen database from which the Smokescreen genotyping array was developed to incorporate recent research into addiction (National Library of Medicine) and into COVID-19 comorbidities (COVID-19 Open Research Dataset, CORD-19). We will use molecular identifiers to select up to 0.175M markers to add to the 1.9M available on the Infinium Global Diversity Array (GDA), currently used by the NIH *All Of Us* Research Program.

Results: Our natural language processing approach of CORD-19 identifies HGNC names and variant identifiers that are weighted based on occurrence. These data will be added to the curated Smokescreen database, and compared to GDA content. A booster for the GDA array will be developed with priorities (existing PRS, novel gene coverage, ethnic coverage, pathway bioinformatics). These addiction and COVID-related markers can be used to prioritize data capture (custom array, querying) and aid in the development of PRS.

Discussion: We envision that applications of this booster will improve our understanding of addiction and the long-term consequences of drug dependence in the time of COVID.