

## Using artificial intelligence for learning substance use disorder biosignatures

James Baurley, BioRealm LLC

There are few biomarkers for substance use disorders (SUDs) and SUD treatments due to their neurobiological complexity. In the quest for new biomarkers, batteries of phenotypic, imaging, digital health, and -omics data are now being collected with an emphasis on prediction and translation. Traditional analysis methods have limitations for discovery or prediction here, due to the dimensionality of the data. The number of variables vastly outnumbers the number of samples (known as  $P \gg N$ ). How can large-scale data be applied to clinical use-cases?

We will introduce machine-learning strategies to handle this problem and help researchers discover new biomarkers and biosignatures for SUDs from high dimensional data. Some of topics addressed will include:

- Concepts and strategies in building these models
- How to apply learned models to other studies (e.g., cohorts and clinical trials)
- How to use the models to predict risk or optimal treatments in individuals
- How to use existing knowledge or evidence
- How to leverage sample and model diversity to improve predictions

Nicotine metabolism, an important biomarker of smoking cessation, will be used to motivate these approaches in practice.