Combating the Drug Epidemic: A Machine-Learning Framework to Predict Novel Drug-Drug Interaction Risks of Illicit Drug Abuse

Wang, Dean (School: Century High School)

Each year, the escalating drug epidemic causes over 100,000 deaths in the US. 80% of these mortalities result from multiple drug use toxicity, which yields unpredictable and potentially fatal drug-drug interactions (DDIs). Thus, it is imperative to discover and prevent DDIs responsible for severe adverse drug reactions (ADRs). Five major drugs—Methamphetamine, Heroin, Cocaine, Oxycodone, and Fentanyl—are responsible for 95% of all overdoses. To discover interactions, I used machine-learning to analyze the ADRs and pharmacological profiles of all 13,117 FDA-approved and 2421 illicit drugs, including their targets, transporters, and enzymes extracted from SIDER and DrugBank. The similarity of drug profiles provides the molecular basis for DDIs, computed with the Jaccard Index. DDI predictions were performed using two machine-learning models: elastic-net logistic regression (GLMnet) and extreme-gradient boosting (XGBoost). In total, I discovered 2746 total interactions, including 841 novel ones. Novel DDIs and severe ADRs were individually found for methamphetamine (novel:199, severe:41), heroin (novel:75, severe:8), cocaine (novel:273, severe:131), oxycodone (novel:116, severe:42), and fentanyl (novel:178, severe:75). GLMnet and XGBoost achieved excellent AUC accuracies at 0.89 (95% CI:0.88-0.90) and 0.93 (95% CI:0.92-0.94), respectively. Further validation was completed with FAERS, a repository of reported ADRs. Overall, this innovative framework predicts novel DDIs and derives a severity spectrum for the world’s five deadliest drugs—critical for combating the drug epidemic. Moreover, this algorithm is readily applicable to other prescription and nonprescription drugs, enhancing drug safety and efficacy.

Awards Won:

The University of Texas at Dallas: Scholarship awards of $5,000 per year, renewable for up to four years