Contact Information: andymytai@alumni.ubc.ca

CHÉOS Centre for Health Evaluation & Outcome Sciences

The Application of Machine Learning and Predictive Modelling to Understand Risk of Overdose (A preliminary model)

Andy Man Yeung Tai, Alireza Kazemi, Amanda Slaunwhite, Kerry Jang, Michael Reinhard Krausz
Department of Psychiatry, Faculty of Medicine, University of British Columbia, Vancouver, Canada

BACKGROUND

- In April 2016, British Columbia declared a public health emergency under the Public Health Act in response to increasing non-fatal and fatal overdose events in the province [1].
- Overdose remains a significant public health challenge. In 2020, 5.5 people died from illicit drug overdoses every day [2].
- Despite previous research identifying risk factors for overdose, there is limited knowledge on their potential to predict fatal opioid overdose events [3].
- In this poster, we will describe the application of these ML methods to create two preliminary models to predict fatal overdose using data from the BC Provincial Overdose Cohort.

OBJECTIVES

- We aimed to create two preliminary models that predicted the risk of fatal overdose and to assess their fit using machine learning techniques

METHODS

- The data utilized from The BC Provincial Overdose Cohort includes many different databases [4, 5, 6, 7]. We utilized data from the OD_Case dataset, BC Coroner's Service and Drug and Poison Information Centre , Mental Health Data Warehouse, British Columbia Emergency Health Services, Client Roster, and lastly Chronic Disease Registries.
- Utilizing machine learning: XGBoost algorithm and the Boruta wrapper algorithm paired with the Random Forest algorithm was utilized to create two preliminary models.
- A confusion control matrix, correlation plot (spearman's correlation) and importance plot were used to determine how the model performed, the weight of each factor and the interactions between each of these factors.

Age	Indicator of prevalence of Acute Myocardial Infarction
Sex	Indicator of prevalence of Acute Myocardial Infarction
Suicide	Indicator of prevalence of Angina
Marital Status	Indicator of prevalence of Asthma
House-hold composition	Indicator of prevalence of Chronic Kidney Disease
Highest level education	Indicator of prevalence of Chronic Obstructive Pulmonary Disease
Current Education	Indicator of prevalence of Depression
Criminal Constraints	Indicator of prevalence of Diabetes Mellitus
Criminal Justice Involvement	Indicator of prevalence of Epilepsy
First age alcohol use	Indicator of prevalence of Heart Failure
First age marijuana use	Indicator of prevalence of Hypertension
First age substance use	Indicator of prevalence of Stroke, Hospitalized
Postal	Indicator of prevalence of Haemorrhagic Stroke
Drug use	Indicator of prevalence of Ischemic Stroke
Drug Class Use	Indicator of prevalence of Stoke, Hospitalized Transient Ischemic Attack
Route of Administration of drug	Indicator of prevalence of Ischemic Heart Disease
Time between the first and last overdose	Indicator of prevalence of Mood and Anxiety Disorders
Total Number of Overdose	Indicator of prevalence of Multiple Sclerosis
City of origin	Indicator of prevalence of Osteoarthritis
Health Authority	Indicator of prevalence of Osteoprosis
Fatal Overdose Case	Indicator of prevalence of Osteoporosis with Fracture
Naloxone Used	Indicator of prevalence of Parkinson's/Parkinsonism
Naloxone Dose	Indicator of prevalence of Rheumatoid Arthritis

Table 1. All Independent Variable titles

If you are interested in these 622 unique classes within these variables, please contact the author.

RESULTS & CONCLUSION

Table 1. Variables to include in a fatal opioid overdose prediction AI tool and their importance

Algorithms	Accuracy	Sensitivity	Specificity	Prevalence	P- Value (Accuracy/ No info Rate)	AUC
XGBoost	98.59	98.35	98.82	0.5	< 0.05	99.93
Boruta Algorithm with Random Forrest	98.43	97.81	99.06	0.5	< 0.05	99.54

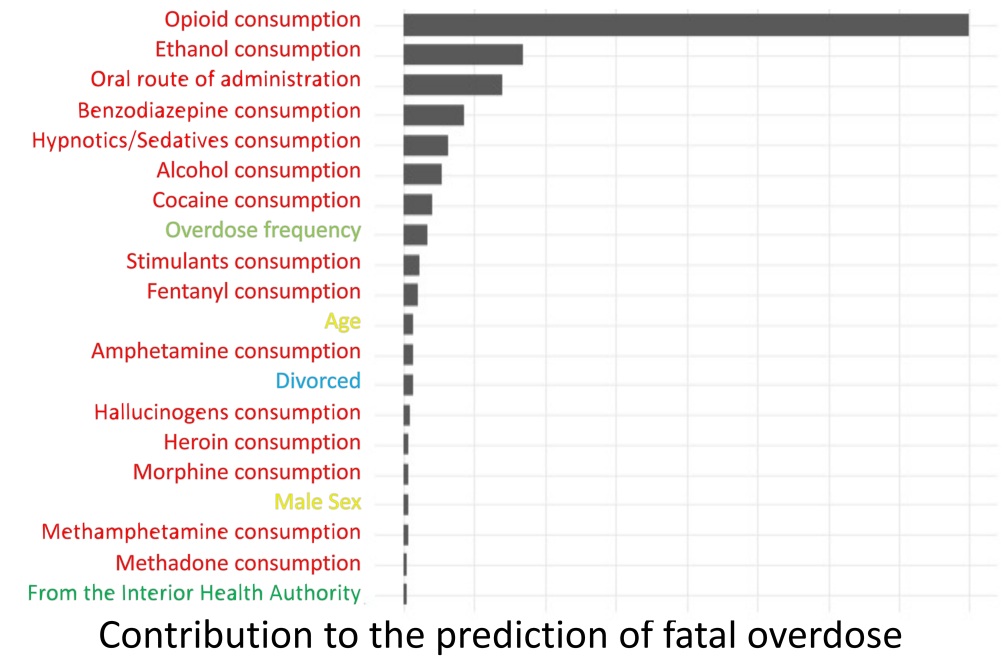

Figure 1. Results from XGBoost

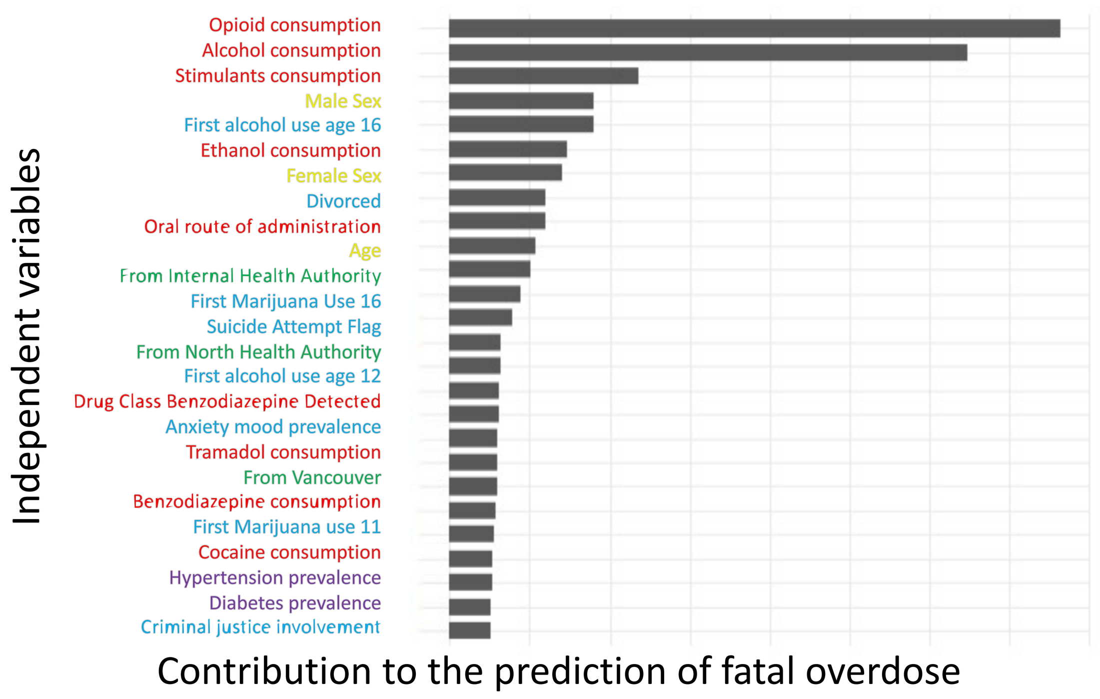

Figure 2. Results from Boruta with RF

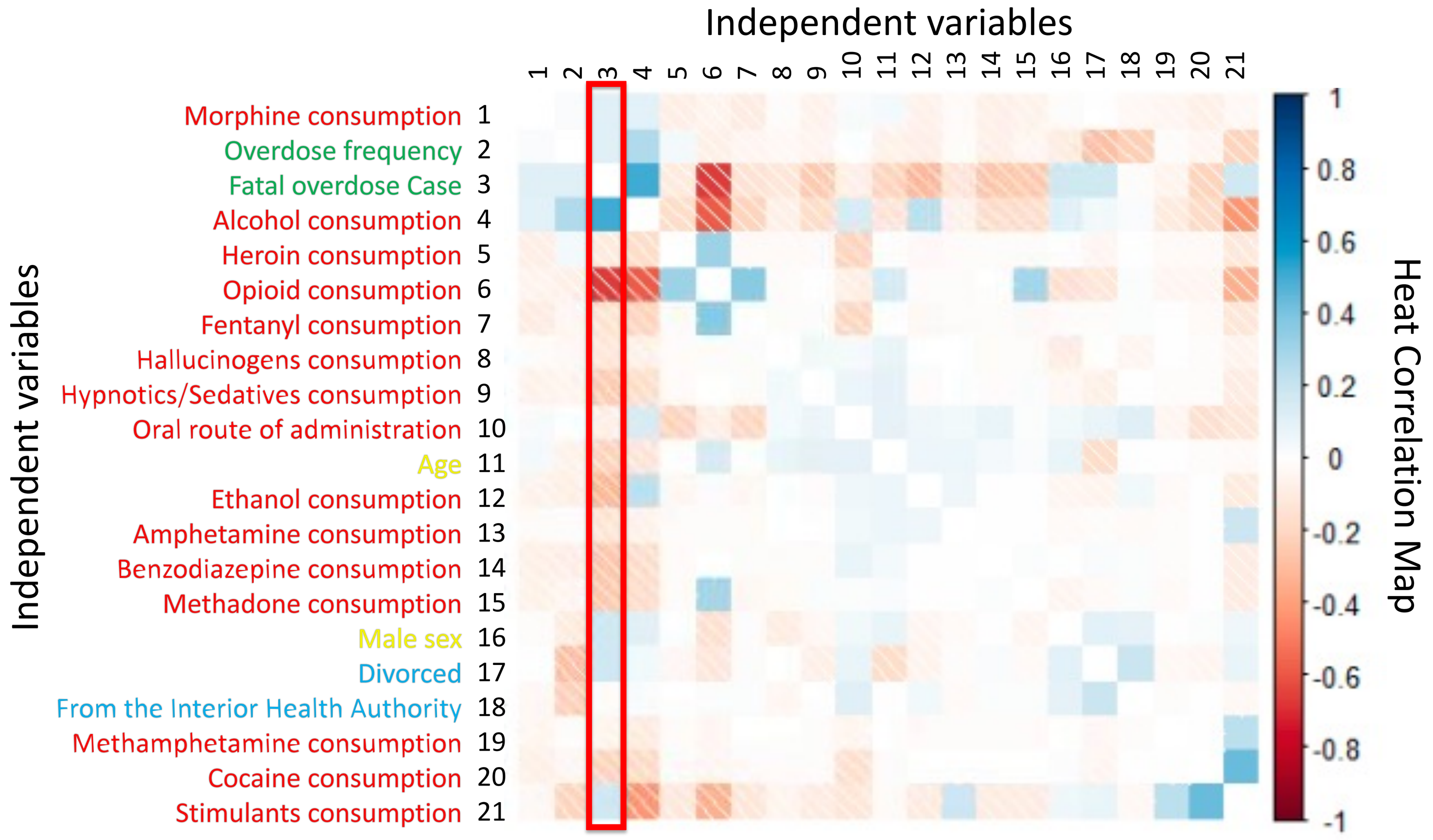

Figure 3. Correlation Plot for XGBoost model (spearman's)

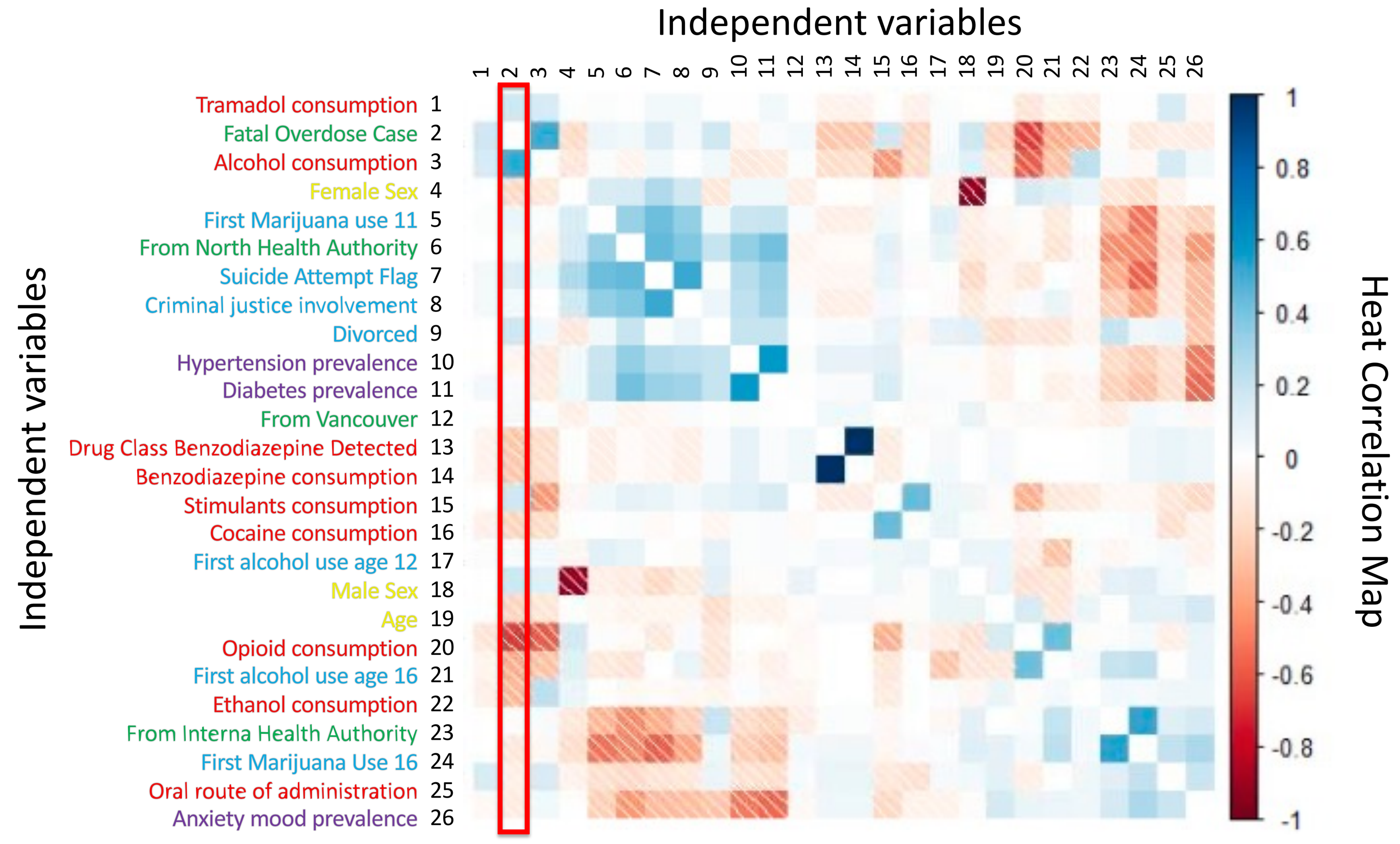

Figure 3. Correlation Plot for Boruta with RF model (spearman's)

- XGBoost model's five most important factors contributing to fatal overdose are the route of administration as inhaled or oral, death by cardiac arrest, the time between recorded overdose and ethanol as a beverage consumed.
- For the Boruta with RF model's five most important factors contributing to fatal overdose are the route of administration as inhaled or oral, drug classes of opioids, stimulants and alcohol.
- In conclusion, the next steps include finding a better class balanced data set that would introduce less bias within the BC Provincial Overdose Cohort.
- This project demonstrates the utility of using ML to understand overdose using existing and derived variables.

All inferences, opinions, and conclusions drawn in this report are those of the authors, and do not reflect the opinions or policies of the Data Steward(s).

References
[1] Special Advisory Committee on the Epidemic of Opioid Overdoses. National Report: Apparent Opioid-Related Deaths in Canada (January 2016 to June 2017) Web-Based Report.; 2017.
[2] "Illicit Drug Toxicity Deaths in BC January 1, 2011–May 31, 2021." BC Centre for Disease Control. 2021 Data Refresh.
[3] "Data Management Framework and Technical Notes." BC Centre for Disease Control: Provincial Overdose Cohort. 2017 Data Refresh.
[4] BC Ministry of Health, Data Stewardship Committee: PharmaNet Data Extract. In. http://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/dataaccess-health-data-central; 2020.
[5] BC Ministry of Health: Consolidation File (MSP Registration & Premium Billing) Data Extract. In.http://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/dataaccess-health-data-central; 2020.
[6] British Columbia Ministry of Health: Vital Statistics Death Events. (Vital Statistics BC). Data Extract. In.http://www2.gov.bc.ca/gov/content/health/conducting-health-researchevaluation/data-access-health-data-central; 2020.
[7] British Columbia Ministry of Health: Medical Services Plan (MSP) Payment Information File. Data Extract In.http://www2.gov.bc.ca/gov/content/health/conducting-health-researchevaluation/data-access-health-data-central; 2020.