



Individualized Prospective Prediction of Opioid Use Disorder

Prédiction prospective individualisée du trouble d'utilisation d'opioïdes

Yang S. Liu, PhD^{1,2} , Lawrence Kiyang, MSc, MPH², Jake Hayward, MD, MPH³ , Yanbo Zhang, MD, PhD¹ , Dan Metes, MSc², Mengzhe Wang, MSc², Lawrence W. Svenson, PhD^{2,4,5,6}, Fernanda Talarico, MSc¹ , Pierre Chue, MD¹, Xin-Min Li, MD, PhD¹, Russell Greiner, PhD^{1,7,8}, Andrew J. Greenshaw, PhD¹ and Bo Cao, PhD^{1,2} 

Abstract

Objective: Opioid use disorder (OUD) is a chronic relapsing disorder with a problematic pattern of opioid use, affecting nearly 27 million people worldwide. Machine learning (ML)-based prediction of OUD may lead to early detection and intervention. However, most ML prediction studies were not based on representative data sources and prospective validations, limiting their potential to predict future new cases. In the current study, we aimed to develop and *prospectively* validate an ML model that could predict individual OUD cases based on representative large-scale health data.

Method: We present an ensemble machine-learning model trained on a cross-linked Canadian administrative health data set from 2014 to 2018 ($n = 699,164$), with validation of model-predicted OUD cases on a hold-out sample from 2014 to 2018 ($n = 174,791$) and prospective prediction of OUD cases on a non-overlapping sample from 2019 ($n = 316,039$). We used administrative records of OUD diagnosis for each subject based on International Classification of Diseases (ICD) codes.

Results: With 6409 OUD cases in 2019 (mean [SD], 45.34 [14.28], 3400 males), our model prospectively predicted OUD cases at a high accuracy (balanced accuracy, 86%, sensitivity, 93%; specificity 79%). In accord with prior findings, the top risk factors for OUD in this model were opioid use indicators and a history of other substance use disorders.

Conclusion: Our study presents an individualized prospective prediction of OUD cases by applying ML to large administrative health datasets. Such prospective predictions based on ML would be essential for potential future clinical applications in the early detection of OUD.

¹Department of Psychiatry, University of Alberta, Edmonton, Alberta, Canada

²Analytics and Performance Reporting Branch, Ministry of Health, Government of Alberta, Edmonton, Alberta, Canada

³Department of Emergency Medicine, University of Alberta, Edmonton, Alberta, Canada

⁴School of Public Health, University of Alberta, Edmonton, Alberta, Canada

⁵Division of Preventive Medicine, University of Alberta, Edmonton, Alberta, Canada

⁶Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada

⁷Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada

⁸Alberta Machine Intelligence Institute (Amii), Edmonton, Alberta, Canada

Corresponding Author:

Bo Cao, Department of Psychiatry Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB T6G 2B7, Canada.

Email: cloudbocao@gmail.com

Abrégé

Objectif: Le trouble d'utilisation d'opioïdes (TULO) est un trouble récidivant chronique avec un modèle problématique d'utilisation d'opioïdes, qui affecte près de 27 millions de personnes dans le monde. La prédiction du TULO basée sur l'apprentissage machine (AM) peut mener à la détection précoce et l'intervention. Toutefois, la plupart des études de prédiction d'AM n'étaient pas basées sur des sources de données représentatives et des validations prospectives, ce qui limitait leur potentiel de prédire de futurs nouveaux cas. Dans la présente étude, nous cherchions à développer et à valider *prospectivement* un modèle d'AM qui pourrait prédire des cas individuels de TULO basés sur des données de santé représentatives à grande échelle.

Méthode: Nous présentons un modèle d'ensemble d'apprentissage machine formé sur des ensembles de données de santé administratives canadiennes croisées de 2014–2018 ($n = 699\ 164$), avec validation de cas de TULO prédits par un modèle dans un contre-échantillon 2014–2018 ($n = 174\ 791$) et une prédiction prospective de cas de TULO sur un échantillon non-chevauchant de 2019 ($n = 316\ 039$). Nous avons utilisé des dossiers administratifs de diagnostics de TULO pour chaque sujet basé sur les codes de la Classification internationale des maladies (CIM).

Résultats: Avec 6 409 cas de TULO en 2019 (moyenne [ET], 45,34 [14,28], 3400 hommes), notre modèle prédisait prospectivement les cas de TULO avec une haute précision (précision équilibrée, 86%, sensibilité, 93%; spécificité 79%). En accord avec les résultats précédents, les principaux facteurs de risque pour le TULO dans ce modèle étaient les indicateurs d'utilisation d'opioïdes et des antécédents de troubles d'utilisation d'autres substances.

Conclusion: Notre étude présente une prédiction prospective individualisée des cas de TULO en appliquant l'AM à de vastes ensembles de données de santé administratives. Ces prédictions prospectives basées sur l'AM seraient essentielles pour les futures applications cliniques potentielles dans la détection précoce des TULO.

Keywords

opioid use disorder, machine learning, administrative health data, electronic health records, prospective validation

Introduction

The opioid crisis continues to worsen in North America.^{1–5} In the United States alone, 10.3 million people reported non-medical opioid use in 2018, and approximately two million were living with opioid use disorder (OUD) that same year.⁶ OUD is a well-established risk factor for opioid overdose and death^{7,8} – for example, in 2018, 128 opioid-related deaths occurred per day, of which 41 were related to prescription opioids.⁹ Between 2016 and 2019, Canada reported more than 16,393 opioid-related deaths, yielding a rate of at least 11 opioid-related deaths per day.¹⁰ Patients with OUD are more likely than others to be hospitalized, visit the emergency department and utilize urgent care services. Once established, OUD is a chronic and disabling disorder associated with a high cost to the healthcare system. In Alberta, Canada, from 2018 to 2019, the annual direct cost to the public health system was \$22,000 for a patient with OUD compared to the average cost per capita of \$2,000–\$3,000 (see Supplementary Methods). If patients *at risk* for OUD could be identified and intervened at the earliest stages of the disease, aversive outcomes are preventable.^{11–16}

Around one-fourth of opioid users will develop OUD,¹⁷ and 8% to 12% of those prescribed opioids for chronic pain will develop OUD¹⁸; therefore, predicting and preventing OUD in this population is pivotal to harm reduction efforts.¹⁹ However, recognizing problematic opioid use in the clinical setting is difficult, especially given that many patients are unwilling or unable to fully communicate their

patterns of prescription and illegal drug use. For this reason, many health systems are turning towards clinical and administrative databases to help identify patients at risk, leveraging data on previously documented health conditions, including substance use disorders, patterns of health-care utilization, socioeconomic status, and history of prescription opioid use. The data available from an administrative database in a single-payer health system provides broad coverage of the population and thus could provide an extensive set of objective indicators. Given the administrative data's strengths and the difficulty collecting and accessing opioid utilization information from illegal sources, we believe that focusing on the OUD in those with prescription opioid use represents the most practical starting point for actionable prediction. While promising, these datasets are highly complex and ever-expanding, creating unique statistical and analytic challenges.

Machine learning (ML)-based modelling has emerged as a fundamental tool for processing health data for patient-level risk estimation due to its capability to tackle complex data and provide individualized predictions.^{20–22} ML is complementary to traditional statistical analysis with its capability of making optimal predictions at the individual level.^{20–22} Compared with conventional statistical analysis, most of which concentrate on group-level significance in the existing data, ML has a built-in process to estimate the model's predictive performance on new data (e.g., cross-validation²³), which ensures generalizability of the model's prediction. ML models

can also make predictions at the individual level in addition to at the group level, which is crucial for real-world applications involving individual outcomes.²² ML models are usually data-driven and able to capture the information and patterns within high-dimensional data through automatic and objective optimization of its parameters and predictive variables, without relying on a-priori hypotheses.²²

Most existing studies on ML for OUD prediction are of limited generalizability due to non-representative sampling. For example, in the United States, only about one-third (34%) of the population is covered by Medicare. Another 9% are uninsured, and the remainder is enrolled in various commercial insurance plans,²⁴ making representative data sampling a major challenge. Also, most existing studies (with the rare exception, see²⁵) use hold-out validation where one portion of the dataset is used for derivation and another for validation, which does not guarantee accurate prediction of future cases.^{26,27} Canada has a centralized, universal and publicly funded health care system,²⁸ where provincially representative data are routinely collected, and nationally representative health administrative data are synthesized and made accessible by Canadian Institute for Health Information. OUD prediction with administrative data has not yet been demonstrated in the Canadian context, leaving a major gap in the path towards cost-effective, targeted prevention efforts to curb the opioid crisis.

In this study, we aimed to develop a novel ML model to prospectively predict individual OUD cases in the future in a large, representative Canadian sample. We also aimed to identify and rank individual risk factors for OUD to improve model interpretability.

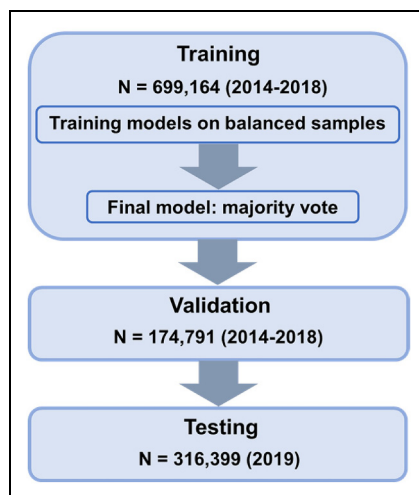


Figure 1. Study flow chart. The ensemble model was built by running a base-learner on 66 different class-balanced subsets of the training set to produce 66 base classifiers; to classify a new patient, the resulting ensemble model runs those 66 classifiers on that patient, then returns the majority vote of their responses, where a case is classified as OUD if over 50% of the models classify the case as OUD. We then tested the ensemble model on the validation and testing set.

Materials and Methods

The study was approved by the University of Alberta ethics review board (Pro00072946). Individual-level information of different types (e.g., demographic, socio-economic, health utilization, etc.) was collected across Provincial government ministries (Alberta, Canada, population of 4.3 million in 2019) and linked and analyzed by the authors (see Supplementary Methods).

Study Cohort

Patients aged 18 years or older who filled an opioid prescription between 2014 and 2019 and with active Alberta Health Insurance coverage were eligible for the study. Patients were enrolled in the year of the most recent opioid prescription fill. We defined potential predicting variables as “features” for ML models. We derived the features using a 5-year window of historical data (e.g., for a patient with the most recent opioid prescription in September 2016, the features were retrieved from September 2011 to September 2016). In line with prior publications on OUD prediction, we developed candidate predictive features based on literature search and human experts curation.^{25,29} We excluded patients receiving cancer medications because high opioid consumption over a short period among cancer patients may represent their routine treatment of pain ($n = 14,079$).³⁰

Patients who developed OUD were identified using the International Classification of Diseases (ICD) codes, including the tenth revision (ICD10) codes: F11.1 (Opioid abuse), F11.2 (Opioid Dependence) available from ambulatory records, and the ninth revision (ICD9) codes: 304.0 (Opioid type dependence), 305.5 (Nondependent opioid abuse), and 304.7 (Combinations of opioid type drug with any other drug dependence) available from physician’s claims.

Supplementary Table 1 shows the 62 features used, including health system utilization indicators (e.g., number of family physician visits), demographics (e.g., age, sex), opioid-specific indicators (e.g., opioid-related poisoning, opioid dosage), other substance abuse, and related disorders (e.g., alcohol, nicotine) and other physical and mental health indicators (e.g., chronic pain, hepatitis, depression).

Sample Derivation

We partitioned the data into three disjoint datasets (see Figure 1): [1] training (retrospective) (2014–2018; $n = 699,164$), [2] validation (retrospective) (2014–18, $n = 174,791$), and [3] testing (prospective) (2019, $n = 316,039$). Because OUD is a chronic condition and we did not have a reliable indicator to identify its termination, prior OUD diagnosis might introduce uncertainty for the OUD status at the enrolment and lead to unreliable classification results. To ensure that the model was trained to classify specific target outcomes (OUD status at the enrollment), the retrospective samples excluded patients with prior records of OUD within 5 years of enrolment ($n = 9669$).

However, during the prospective testing, we kept patients with prior history of OUD to verify the generality of the model predictions. Patients included in the prospective sample were excluded from the retrospective samples ($n = 189,400$). The retrospective sample was split into training and validation samples based on a 4 to 1 ratio, stratified to match the distribution of OUD, Age and Sex. This stratification was to reduce validation variance and to ensure OUD, Age and Sex were equally represented in the training and validation datasets.

Class imbalance is a common issue when building a model to predict rare events.³¹ If an ML algorithm was trained on one instance of a minority class but many examples of a majority class, this may lead to poor classification accuracy for the minority class. To enhance model performance in this setting, undersampling of the majority class is a commonly used technique.³² In addition, ensemble methods can improve prediction performance³³ by using a set of classifiers to classify new data points through prediction voting. We adopt an ensemble method based on results voting from classifiers trained on multiple class-balanced samples to address class imbalance issues while optimizing predictive efficacy. The ratio of OUD to non-OUD cases was approximately 1 to 66 (1.49% OUD), for the training data, we randomly split the majority class (non-OUD) into 66 equal-sized samples (each $n = 12,860$) matched to the size of the minority class (OUD). Each non-OUD sample was combined with the OUD sample to create 66 class-balanced training samples.

Data Preparation

SAS 9.4 was used for extracting and preparing the data for modelling. Python 3.6 with scikit-learn 0.22.1 package was used for running predictive models. The majority of the features were treated as continuous, primarily measured in *number of years* (frequency of occurrences), except the binary features – Sex (Male, Female), Family Status (Couple/Single), Long-Term Use (LTU) indicator (yes/no),³⁴ Receiving Government Financial Aid (yes/no) – and non-binary categorical features: Family Income (Low, Middle, High) and Opioid Dosage (Low, Moderate, High).³⁴ We dummy coded non-binary categorical features. Table 1 summarizes the demographic information. Variable derivations are described in Supplementary Table 1. Because our data are from a single-payer system where all interactions with the health system were recorded and derived based on multiyear aggregated frequency, no record was coded as 0 and interpreted as no occurrence. The stratification of OUD, Age and Sex for training and validation datasets was performed using the StratifiedKfold function. All continuous features were scaled using RobustScaler from the scikit-learn package. Compared with mean and unit variance scaling, RobustScaler is a method of data standardization that is more resistant to adverse impact of outliers, which scales the data based on the median and an interquartile range between the 1st and 3rd quartile. The

testing data from 2019 were transformed by applying the same scale used on the 2014–18 training sample.

Modelling

We developed an ensemble classification method by running logistic ridge regressions (L_2 regularization) classifiers on each of the 66 training samples to learn base models for classifying OUD (1 = OUD, 0 = non-OUD), optimizing for balanced accuracy, defined as the average of sensitivity and specificity. We used internal five-fold cross-validation to select the best parameter value for regularization. Then, we ran each base classifier on the validation set (Figure 1). All features were ranked for each classifier based on the absolute value of the coefficients, and an averaged rank was obtained to determine the top contributing features. This ensemble model predicted OUD for a patient if more than 50% of the base classifiers classified that patient's label as OUD.

Results

Model Performance

The ensemble model reached a balanced accuracy of 91.5% in the validation data, with a sensitivity of 87.9% and specificity of 95.1%. For the testing data, the model reached a balanced accuracy of 86.0%, a sensitivity of 93.0%, and a specificity of 78.9%, with an Area Under the Curve (AUC) of 0.94. The sensitivity and specificity trade-offs in the 2019 testing cohort were illustrated in the receiver operating characteristic (ROC) curve and confusion matrix plot displayed in Figure 2.

Predictive Risk Factors

The ten top-ranked predictors (see Table 2 for case definitions) were dominated by opioid use indicators and other substance abuse disorders (Table 3). Predictors measured in the *number of years* were highly skewed, with the majority being 0 or 1 year. For presentation clarity, *number of years* was collapsed into two categories (<1 or ≥ 1 years; see Supplementary Figure 1 for complete distributions). The most substantial risk factors for OUD were Opioid-Related Poisoning, LTU indicator and High Dose, representing 22.64%, 67.43% and 5.90% of the OUD sample, respectively). Other substance abuse disorders were also the top-ranked predictors, including Sedative Hypnotic Related Disorders, Polysubstance Related Disorders, Amphetamine Related Disorders, Cocaine Related Disorders, Cannabis-Related Disorders and Hallucinogen Related Disorders. Low Dose of opioids was the only protective factor among the top 10 predictors, where 97.53% of non-OUD patients had ≥ 1 year of low dose (See the Low Opioid Dosage row in Table 3).

Table 1. Cohort Demographics.

Features	Measurements	Training and Validation (2014–2018)		Testing (2019)	
		OUD (n = 12,860)	Non-OUD (n = 861,095)	OUD (n = 6409)	Non-OUD (n = 309,990)
Age	Age at enrolment (mean ± standard deviation)	43.79 ± 14.67	46.69 ± 18.55	45.34 ± 14.28	51.56 ± 18.27
Sex	Male	55.61%	49.35%	53.05%	49.92%
	Female	44.39%	50.65%	46.95%	50.08%

Discussion

In the current study, we developed and *prospectively* validated an ML model that could predict individual OUD cases in the following year based on representative large-scale health data. Performance of our model on hold-out validation data reached an overall balanced accuracy of 91.5%, and performance on prospective testing data achieved an overall balanced accuracy of 86.0% with an AUC of 0.94. We found that opioid utilization and other substance use disorders are the predominant risk factors driving our model predictions, indicating an increased risk for those with long-term heavy use of prescription opioids and a history of other substance use disorders.

Predictive Performance

Prospective prediction of OUD is challenging, given the ongoing significant and rapid shifts in opioid-related policy, culture and medical management.²⁷ However, our prospective accuracy is comparable to or better than studies using hold-out validation and a similar case definition for OUD. Other studies included additional features such as lab tests and vital signs,³⁵ additional diagnostic phenotypes developed based on unsupervised learning,³⁶ and a larger quantity and broader spectrum of demographic and clinical features,^{37–39} as well as utilizing advanced ML models such as decision trees,^{37–39} gradient boosted trees⁴⁰ and deep neural network learning.^{39–41} Our findings demonstrate that a ML model based on retrospective data can provide accurate predictions of individual OUD cases in a prospective sample.

The lower accuracy of prospective predictions (86%) compared to the accuracy of validation on 2018 data (91.5%) may reflect the lag time between developing OUD and receiving a clinical diagnosis, i.e., a portion of false positives could represent patients with OUD who are not yet diagnosed. Therefore, false-positive cases may reflect a high-risk group in the early stages of OUD and provide an important opportunity for clinical and policy-level interventions. These cases will need to be further followed up for future investigations.

It is difficult to compare our results with some of the previous studies because existing literature on ML for opioid-related risk prediction is inconsistent and includes a varied array of models that predict similar but fundamentally different outcomes such as OUD,^{25,35,36,40–43} opioid misuse,^{38,44–46} and opioid overdose

(See Tseregounis & Henry, 2021, for a review⁴⁷). Opioid misuse refers to the non-medical use of prescription opioids, for example, for euphoric effect. Opioid misuse increases the risk of opioid-related overdose and development of OUD⁴⁸ but is not associated with a unique diagnostic code and is therefore difficult to identify and predict using administrative data. Cases referred to as opioid misuse in previous studies may include OUD and other opioid-related adverse outcomes. Still, the prediction of misuse cases may be difficult to interpret from the clinical perspective and to be compared with our study.

Risk Factors of OUD

In a clinical setting, risk factors for OUD include past or current substance abuse, diagnoses of psychiatric disorders, younger age, male sex, and social or family environment that encourage misuse.^{49–51} Our predictive model identified that the highest risk factors for OUD were tied to opioid use indicators and other substance use disorders, consistent with the literature.^{49,51} Given that ‘opioid-related poisoning’ is considered synonymous with opioid overdose, it is perhaps not surprising that *opioid-related poisoning* emerged as the strongest predictor of OUD. However, post-hoc analysis showed that opioid-related poisoning and OUD have a low linear correlation in the 2019 testing data ($r = 0.298$), with only 22.64% of OUD patients experiencing one or more events of opioid-related poisoning. This is consistent with existing evidence that opioid overdoses infrequently occur in prescription opioid users,^{2,7} and reducing opioid prescription availability may lead to higher overdose rates.⁵² Of the non-opioid substance use disorders examined, polysubstance and amphetamine-related disorders were the strongest predictors of OUD, present in 41.25% and 22.84% of the 2019 sample, respectively. A range of other non-opioid substance use disorders, including sedative-hypnotics, cannabis, cocaine and hallucinogens, also predicted OUD, albeit to a lesser extent (present in fewer than 10% of OUD patients). Two-thirds of the OUD population (67.34%) experienced LTU during the year of enrolment, compared to 19% of non-OUD patients. The distribution of top predictors of OUD and non-OUD patients suggests that each of the individual factors is not powerful enough to predict OUD alone. The OUD prediction needs to harness the combined effects of predictors through the power of multivariate analysis.

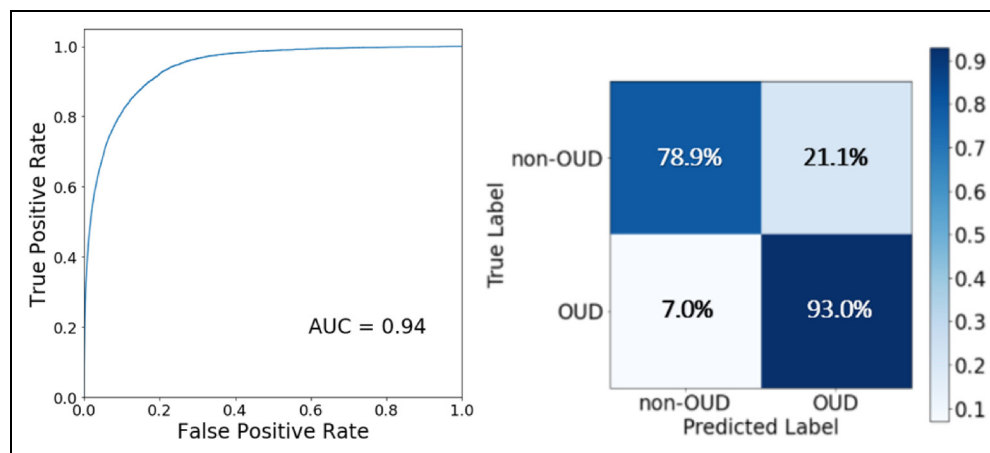


Figure 2. Receiver operating characteristics curve and confusion matrix. Receiver operating characteristics curve and confusion matrix for prediction of OUD in the testing cohort for OUD cases. True Positive Rate is the rate model predicts a positive OUD label correctly, equivalent to Sensitivity, whereas False Positive Rate is the rate model predicts a positive OUD label incorrectly, equivalent to 1 – Specificity. AUC stands for area under the curve.

Table 2. Feature Case Definitions.

Features	Case Definition
Opioid-related poisoning	ICD9 code 965.0 and/or ICD 10 codes T40.0, T40.1, T40.2, T40.3, T40.4, T40.6, occurring within a five-year window prior to enrollment year and measured in <i>number of years</i> (frequency ranges from 0 to 5 years)
Long-term use indicator	Dispensation duration > 120 days, or dispensation frequency > 10 or a > 90 days period of continuous opioid prescribing following the date of enrollment ³⁴
High/medium/low dosage	Daily morphine equivalent doses of > 200 mg, <200 mg and > 90 mg, or < 90 mg, respectively, during the enrollment year ³⁴

Note. Substance-related disorders including Sedative and Hypnotic, Polysubstance, Amphetamine, Cocaine, Cannabis, and Hallucinogen were all defined like Opioid Related Poisoning, by corresponding ICD9 and ICD10 codes and measured in *number of years*. For more details on all features, see Supplementary Table 1 in Supplementary Materials.

Implications

Prospective prediction based on historical data is a significant step towards deploying OUD prediction and risk assessment programs at the population level in real time, arguably a new gold standard for applied precision medicine. Early identification of patients at risk for developing OUD might enable preventative and harm reduction interventions that mitigate harms and allow cost savings at the system level.⁵³ The utility and efficacy of predictive algorithm on OUD have already been studied in the context of primary health care, with longitudinal sample and prospective validation.⁵⁴⁻⁵⁶ Algorithm-guided physicians benefited from clinically actionable insights and were able to prevent

OUD by limiting initial opioid prescription and improving patient’s pain treatment outcomes.⁵⁵ It is worth noting that the validated clinical OUD detection tool utilized similar patient history information reported in our study (e.g., illegal drug use, history of prescription drug use, mental disorders, age) along with patient’s genomic testing data (AUC = 0.83). This suggests that the addition of administrative health-based predictions with ML may further enhance the screening efficacy of existing prediction tools used clinically and will yield superior benefits to enhance patient outcomes with evidence of higher predictive performance (e.g., our study had an AUC of 0.92). The current study can be viewed as a proof of concept for future applications of administrative health data on OUD prediction and prevention but not a replacement for clinical judgment and diagnoses. More research is needed to further optimize the methods and determine an optimal strategy for real-world adaptation and translation. One potential route is to leverage existing healthcare resources and incorporate the prediction result as a risk indicator within the electronic medical record system. When a patient is flagged as high risk for developing OUD, a clinician might then be encouraged to further screen for modifiable OUD risks and alter future opioid prescribing strategies. The risk factors identified by the model may further assist clinicians to go through the medical history of patients, while save their time. Increasingly, patients have access to their own electronic medical records, presenting novel opportunities for encouraging a better understanding of personal health risks, including risks of OUD. Such insights might become particularly important, e.g., in the context of pain treatment allowing patients to make risk-informed decisions about available options on the use of prescription opioids. Thus, innovative and interactive patient/provider interface might translate complex risk models derived from highly dimensional historical health data into actionable insights for effective prevention, diagnosis and treatment of OUD.

Table 3. Summary of Top 10 Predictors in the Testing Data.

Predictors	OUD Risk	Non-OUD (n = 309,990)		OUD (n = 6409)		χ^2	r_ϕ	p
		< 1 year	≥ 1 year	< 1 year	≥ 1 year			
Opioid-related poisoning	Increase	99.04%	0.96%	77.36%	22.64%	212,393	.82	<0.001
Sedative hypnotic-related disorders	Increase	99.76%	0.24%	93.45%	6.55%	6822	.15	<0.001
Polysubstance-related disorders	Increase	97.56%	2.44%	58.75%	41.25%	30,293	.31	<0.001
Long-term use indicator	Increase	81.06%	18.94%	32.66%	67.34%	9221	.17	<0.001
Amphetamine-related disorders	Increase	99.17%	0.83%	77.16%	22.84%	24,150	.28	<0.001
Cocaine-related disorders	Increase	99.23%	0.77%	87.64%	12.36%	8480	.16	<0.001
High opioid dosage	Increase	99.55%	0.45%	94.10%	5.90%	3347	.10	<0.001
Cannabis-related disorders	Increase	98.43%	1.57%	90.72%	9.28%	2200	.08	<0.001
Low opioid dosage	Decrease	2.47%	97.53%	14.98%	85.02%	3709	.11	<0.001
Hallucinogen-related disorders	Increase	99.63%	0.37%	98.36%	1.64%	257	.03	<0.001

Note. The predictors were presented in the order of ranked importance. The percentage values were calculated based on the proportion of < 1 year and ≥ 1 year in patients' subgroups based on the OUD label, two-way χ^2 , r_ϕ and p values were calculated based on raw counts.

Limitations

The results of our study need to be interpreted with caution due to their limitations. Firstly, while our ML-based approach optimizes accuracy of individual-level predictions, model interpretability could be improved with more traditional statistical approaches and further validation. Relatedly, the study data represents the health care system users in Alberta, Canada; thus, generalizability to other health care systems may be limited. Secondly, some elements in our administrative data were designed for managing physician payments (i.e., billing data) and may be suboptimal for capturing clinical events, potentially leading to missed OUD diagnoses,⁵⁷ or conversely, false OUD labels in cases of suspected or unconfirmed diagnosis. The same limitation applies to the candidate risk factors, while no data is available for those do not utilize the healthcare system, e.g., patients obtain substances through illegal sources, or recently moved to Alberta. Our data also lack a clear indicator for treated OUD or OUD in remission. An alternative solution to validate OUD labels is to collect clinical diagnostic information from the clinical setting (e.g., from electronic medical records) and make the data available to administrative health data custodians.

Further, our OUD definition relied on the premise that OUD develops after LTU and, therefore, LTU during the year of enrolment acted as a predictor of OUD. It is possible that in rare cases, OUD developed before LTU criteria were satisfied. However, the administrative data may have latencies for relative event records, which makes it challenging to consolidate the exact temporal sequence of the events. Future studies should better differentiate OUD in remission from active OUD and delineate the temporal relationship between LTU and OUD.

In addition, given that interpretability of the model is important for government policymaking, we chose ridge regression to emphasize interpretability at the cost of

potentially inferior predictive performance compared to non-linear models. Multiple algorithms should be compared for prospective prediction performance in future studies where the context requires real-life applications. Finally, future studies should explore the optimal method to address data shift,²⁷ and to continuously update the model to maintain high predictive accuracy.

Conclusion

Our study developed and prospectively validated a promising novel approach that could predict individual OUD cases based on ML and representative large-scale health data. Our findings suggest that early detection of OUD is possible with such a data-driven approach, which may enable timely clinical intervention and policy changes relevant to OUD and other opioid-related outcomes to help curb the opioid crisis.

Acknowledgement

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program, Alberta Innovates, Mental Health Foundation, MITACS Accelerate program, NARSAD Young Investigator Grant of The Brain & Behavior Research Foundation, Simon & Martina Sochatsky Fund for Mental Health, the Alberta Synergies in Alzheimer's and Related Disorders (SynAD) program and University of Alberta Hospital Foundation. The funding sources had no impact on the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication.

Authors' Note

This paper is dedicated to the memory of our dear mentor and colleague, Larry Svenson, who passed away while we prepared the revision.

Author Contributions

Yang S. Liu designed the study, processed, quality-controlled and analyzed the data, drafted the manuscript, and critically edited the draft of the manuscript. Lawrence Kiyang was involved in designing the study, processed, quality-controlled data analysis and critically revised the manuscript draft. Jake Hayward and Yanbo Zhang critically edited the draft of the manuscript. Dan Metes was involved in designing the study, quality-controlled and analyzed the data and edited the manuscript draft. Mengzhe Wang is involved in designing the study, quality-controlled and analyzing the data, and editing the manuscript draft. Lawrence W. Svenson provide instrumental support, revised the draft of the manuscript. Fernanda Talarico is involved in editing the figures and the draft of the manuscript. Pierre Chue, Xin-Min Li and Russell Greiner edited the draft of the manuscript. Andrew J. Greenshaw provided instrumental support and critically edited the draft of the manuscript. Bo Cao designed the study, supervised the data processing, provided instrumental support, revised the draft of the manuscript and supervised the paper writing. Yang S. Liu and Bo Cao had full access to all the data in the study and took responsibility for the data's integrity and the accuracy of the data analysis.






Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported by the Brain and Behavior Research Foundation NARSAD Young Investigator Grant, Alberta Innovates, Mitacs, Alberta Synergies in Alzheimer's and Related Disorders (SynAD) program, Simon & Martina Sochatsky Fund for Mental Health, University of Alberta Hospital Foundation, Canada Research Chairs program, Mental Health Foundation.

ORCID iDs

Yang S. Liu  <https://orcid.org/0000-0003-0406-8056>
 Jake Hayward  <https://orcid.org/0000-0003-4919-8946>
 Yanbo Zhang  <https://orcid.org/0000-0002-2421-157X>
 Fernanda Talarico  <https://orcid.org/0000-0001-6114-4233>
 Bo Cao  <https://orcid.org/0000-0001-9338-3271>

Supplemental Material

Supplemental material for this article is available online.

References

- Fischer B, Pang M, Tyndall M. The opioid death crisis in Canada: crucial lessons for public health. *Lancet Public Health*. 2019;4(2):e81-e82.
- Rudd RA. Increases in drug and opioid-involved overdose deaths—United States, 2010–2015. *Morb Mortal Wkly Rep*. 2016;65.
- Strang J, Volkow ND, Degenhardt L, et al. Opioid use disorder. *Nat Rev Dis Primers*. 2020;6(1):1-28.
- Blanco C, Volkow ND. Management of opioid use disorder in the USA: present status and future directions. *Lancet*. 2019;393(10182):1760-1772.
- Vojtila L, Pang M, Goldman B, et al. Non-medical opioid use, harms, and interventions in Canada – a 10-year update on an unprecedented substance use-related public health crisis. *Drugs Educ Prev Pol*. 2020;27(2):118-122.
- Substance Abuse and Mental Health Services Administration (SAMHSA). Key substance use and mental health indicators in the United States: Results from the 2018 National Survey on Drug Use and Health. HHS Publication No PEP19-5068, NSDUH Series H-54. Epub ahead of print 2019. DOI: 10.1016/j.drugalcdep.2016.10.042.
- Smolina K, Crabtree A, Chong M, et al. Patterns and history of prescription drug use among opioid-related drug overdose cases in British Columbia, Canada, 2015–2016. *Drug Alcohol Depend*. 2019;194:151-158.
- Fischer B, Jones W, Varatharajan T, et al. Correlations between population-levels of prescription opioid dispensing and related deaths in Ontario (Canada), 2005–2016. *Prev Med*. 2018;116:112-118.
- National Institute on Drug Abuse. Overdose Death Rates. National Institute on Drug Abuse; 2021 [accessed July 20, 2021]. Available from: <https://www.drugabuse.gov/drug-topics/trends-statistics/overdose-death-rates>.
- Public Health Agency of Canada. Opioid-related harms in Canada. 2020 [accessed August 27, 2020]. Available from: <https://health-infobase.canada.ca/substance-related-harms/opioids/>.
- Chang HY, Kharrazi H, Bodycombe D, et al. Healthcare costs and utilization associated with high-risk prescription opioid use: a retrospective cohort study. *BMC Med*. 2018;16(69). DOI: 10.1186/s12916-018-1058-y.
- Florence CS, Zhou C, Luo F, et al. The economic burden of prescription opioid overdose, abuse, and dependence in the United States, 2013. *Med Care*. 2016;54(10):901-906.
- Meyer R, Patel AM, Rattana SK, et al. Prescription opioid abuse: a literature review of the clinical and economic burden in the United States. *Popul Health Manag*. 2014;17(6):372-387.
- Rice JB, Kirson NY, Shei A, et al. Estimating the costs of opioid abuse and dependence from an employer perspective: a retrospective analysis using administrative claims data. *Appl Health Econ Health Policy*. 2014;12(4):435-446.
- Roland CL, Joshi AV, Mardekian J, et al. Prevalence and cost of diagnosed opioid abuse in a privately insured population in the United States. *J Opioid Manag*. 2013;9(3):161-175.
- Jones W, Kurdyak P, Fischer B. Examining correlations between opioid dispensing and opioid-related hospitalizations in Canada, 2007–2016. *BMC Health Serv Res*. 2020;20(1):677.
- Schlag AK. Percentages of problem drug use and their implications for policy making: a review of the literature. *Drug Sci Pol Law*. 2020;6:1-9.
- Vowles KE, McEntee ML, Julnes PS, et al. Rates of opioid misuse, abuse, and addiction in chronic pain: a systematic review and data synthesis. *Pain*. 2015;156(4):569-576.

19. Degenhardt L, Grebely J, Stone J, et al. Global patterns of opioid use and dependence: harms to populations, interventions, and future action. *Lancet*. 2019;394(10208):1560-1579.
20. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA*. 2018;319(13):1317-1318.
21. Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. *J Glob Health*. 2018;8(2):020303.
22. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods*. 2018;15(4):233-234.
23. Browne MW. Cross-Validation methods. *J Math Psychol*. 2000;44(1):108-132.
24. Berchick ER, Hood E, Barnett JC. Health insurance coverage in the United States: 2017. Current population reports. Washington DC: US Government Printing Office.
25. Gao W, Leighton C, Chen Y, et al. Predicting opioid use disorder and associated risk factors in a Medicaid managed care population. *Am J Manag Care*. 2021;27(4):148-154.
26. Quionero-Candela J, Sugiyama M, Schwaighofer A, et al. Dataset shift in machine learning. Cambridge, MA: The MIT Press; 2009. p. 248.
27. Subbaswamy A, Saria S. From development to deployment: dataset shift, causality, and shift-stable models in health AI. *Biostatistics*. 2020;21(2):345-352.
28. Canadian Institute for Health Information. Data Holdings. CIHI; 2021 [accessed September 9, 2021]. Available from: <https://www.cihi.ca/en/access-data-and-reports/make-a-data-request/data-holdings>.
29. Lo-Ciganic W-H, Huang JL, Zhang HH, et al. Evaluation of machine-learning algorithms for predicting opioid overdose risk among Medicare beneficiaries with opioid prescriptions. *JAMA Netw Open*. 2019;2(3):e190968-e190968.
30. Cancer Drugs & DINs List. ODANO. 2018 [accessed July 24, 2020]. Available from: <https://odano.ca/cancer-drugs-dins-list/>.
31. Guo X, Yin Y, Dong C, et al. On the class imbalance problem. In: 2008 Fourth international conference on natural computation. Berlin, Heidelberg: Springer; 2008. P.192-201.
32. Sun Y, Wong AKC, Kamel MS. Classification of imbalanced data: a review. *Int J Patt Recogn Artif Intell*. 2009;23(04):687-719.
33. Dietterich TG. Ensemble methods in machine learning. In: International workshop on multiple classifier systems. Springer; 2000. P.1-15.
34. Von Korff M, Saunders K, Ray GT, et al. Defacto long-term opioid therapy for non-cancer pain. *Clin J Pain*. 2008;24(6):521-527.
35. Ellis RJ, Wang Z, Genes N, et al. Predicting opioid dependence from electronic health records with machine learning. *BioData Min*. 2019;12(1). 2019;12:3. DOI: 10.1186/s13040-019-0193-0.
36. Gong JJ, Jacobs AZ, Stuart TE, et al. Discovering heterogeneous subpopulations for fine-grained analysis of opioid use and opioid use disorders. arXiv preprint {arXiv}:1811; 2018.P.04344.
37. Hastings JS, Howison M, Inman SE. Predicting high-risk opioid prescriptions before they are given. *Proc Natl Acad Sci U S A*. 2020;171(4):1917-1923. DOI: 10.1073/pnas.1905355117.
38. Hasan MM, Patel MR, Modestino AS, et al. A Novel Big Data Analytics Framework to Predict the Risk of Opioid Use Disorder. arXiv preprint {arXiv}:1904; 2019. P.03524.
39. Lo-Ciganic W-H, Huang JL, Zhang HH, et al. Using machine learning to predict risk of incident opioid use disorder among fee-for-service Medicare beneficiaries: a prognostic study. *PLoS one*. 2020;15(7):e0235981.
40. Segal Z, Radinsky K, Elad G, et al. Development of a machine learning algorithm for early detection of opioid use disorder. *Pharmacol Res Perspect*. 2020;8(6):e00669.
41. Che Z, St Sauver J, Liu H, et al. Deep learning solutions for classifying patients on opioid use. In: AMIA Annual Symposium Proceedings, Washington, DC. Vol. 2017, P.525.
42. McDougall S, Annareddy P, Madiraju P, et al. Predicting Opioid Overdose Readmission and Opioid Use Disorder with Machine Learning. In: 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA; 2020, P.4892-4901.
43. Wadekar AS. Understanding opioid use disorder (OUD) using tree-based classifiers. *Drug Alcohol Depend*. 2020;208:107839-107839.
44. Afshar M, Sharma B, Bhalla S, et al. External validation of an opioid misuse machine learning classifier in hospitalized adult patients. *Addict Sci Clin Pract*. 2021;16(1):1-11.
45. Dufour R, Mardekian J, Pasquale MK, et al. Understanding predictors of opioid abuse: predictive model development and validation. *Am J Pharm*. 2014;6(5):208-216.
46. Han D-H, Lee S, Seo D-C. Using machine learning to predict opioid misuse among U.S. adolescents. *Prev Med*. 2020;130:105886-105886.
47. Tsergounis IE, Henry SG. Assessing opioid overdose risk: a review of clinical prediction models utilizing patient-level data. *Transl Res*. 2021;234:74-87.
48. Volkow ND, Jones EB, Einstein EB, et al. Prevention and treatment of opioid misuse and addiction: a review. *JAMA Psychiatry*. 2019;76(2):208-216.
49. Cragg A, Hau JP, Woo SA, et al. Risk factors for misuse of prescribed opioids: a systematic review and meta-analysis. *Ann Emerg Med*; 2019. 2019;74(5):634-646. DOI: 10.1016/j.annemergmed.2019.04.019.
50. Webster LR. Risk factors for opioid-use disorder and overdose. *Anesth Analg*. 2017;125(5):1741-1748.
51. McAnally HB. Opioid dependence risk factors and risk assessment. In: Opioid dependence. Cham, Switzerland: Springer International Publishing; 2018. P.233-264.
52. Fischer B, Jones W, Tyndall M, et al. Correlations between opioid mortality increases related to illicit/synthetic opioids and reductions of medical opioid dispensing - exploratory analyses from Canada. *BMC Public Health*. 2020;20(1):143.
53. Bharel M, Bernson D, Averbach A. Using data to guide action in response to the public health crisis of opioid overdoses. *NEJM Catal Innov Care Delivery*. 2020;1(5).

54. Lee C, Sharma M, Kantorovich S, et al. A predictive algorithm to detect opioid use disorder: what is the utility in a primary care setting? *Health Serv Res Manag Epidemiol.* 2018;5:1-8.
55. Brenton A, Lee C, Lewis K, et al. A prospective, longitudinal study to evaluate the clinical utility of a predictive algorithm that detects risk of opioid use disorder. *J Pain Res.* 2018;11:119-131.
56. Sharma M, Lee C, Kantorovich S, et al. Validation study of a predictive algorithm to evaluate opioid use disorder in a primary care setting. *Health Serv Res Manag Epidemiol.* 2017;4:1-9.
57. Hylan TR, Korff MV, Saunders K, et al. Automated prediction of risk for problem opioid use in a primary care setting. *J Pain.* 2015;16(4):380-387.