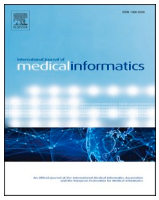
International Journal of Medical Informatics 203 (2025) 106031



Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf





A web-based tool utilizing machine learning algorithms for predicting illicit drug use in emergency departments

Tsung-Chien Lu [a,b,1], Chih-Chuan Lin [c,d,1], Te-I Weng [a,e,f], Fan-Ya Chou [a,b], Cheng- Chung Fang [a,b,e,*], TEDAS Research Group[2]

[a] Department of Emergency Medicine, National Taiwan University Hospital, Taipei City, Taiwan
[b] Department of Emergency Medicine, College of Medicine, National Taiwan University, Taipei City, Taiwan
[c] Department of Emergency Medicine, Linkou Chang Gung Memorial Hospital and College of Medicine, Chang Gung University, Taoyuan City, Taiwan
[d] School of Medicine, Chang-Gung University, Taoyuan City, Taiwan
[e] Forensic and Clinical Toxicology Center, College of Medicine and National Taiwan University Hospital, National Taiwan University, Taipei City, Taiwan
[f] Department of Forensic Medicine, College of Medicine, National Taiwan University, Taipei City, Taiwan

ARTICLE INFO

Keywords:
Machine learning
Emergency department
Illicit drug use
Novel psychoactive substances

ABSTRACT

Background: Identifying illicit drug use through urine testing is time-consuming in the era of new psychoactive substances. This study aimed to develop a machine learning (ML) prediction model for early identification of illicit drug use in suspected emergency department (ED) patients.
Methods: Data from the Taiwan Emergency Department Drug Abuse Surveillance (TEDAS) database (2020–2023) was used. Six feature categories—demographics, triage data, referral source, symptoms, physical findings, and clinical characteristics—were included. The primary outcome was positive urine results for illicit drugs, confirmed by liquid chromatography-tandem mass spectrometry. Data were divided chronologically into training/validation and testing sets. Three supervised ML algorithms, including random forest, CatBoost, and light gradient boosting machine, were tested using K-fold cross-validation, and performance was evaluated by the area under the receiver operating characteristic curve (AUC) in the test set.
Results: The analysis included 13,615 urine test results from ED cases, identifying 3,185 positive cases (23.4%). A total of 9,529 cases (2020–2022) formed the training/validation cohort, and 4,086 (2023) were used for testing. Twenty features were used to construct the prediction model. The CatBoost classifier performed best, achieving an AUC of 0.846 (95% confidence interval [CI]: 0.831–0.859) in the testing cohort. A web-based tool and mobile apps were implemented to assist emergency physicians in predicting illicit drug use.
Conclusions: The machine learning model effectively predicts illicit drug use in ED patients and has been successfully implemented for free access. Further analysis is needed to assess post-implementation performance and its potential for use in other countries.

1. Introduction

The global surge of novel psychoactive substances (NPS), including synthetic cathinones, synthetic cannabinoids, phenethylamines, piperazines, and tryptamines, poses a significant challenge to public health and substance abuse management [1,2,3]. Marketed as "research chemicals," "bath salts," or "plant foods", NPS serve as legal alternatives to traditional illicit drugs like methamphetamine, cocaine, and 3,4-methylenedioxymethamphetamine (MDMA), contributing to their widespread use [4]. Polysubstance use, where patients consume multiple drugs simultaneously, has become increasingly prevalent, with study like Sweden's STRIDA project reporting that approximately 50 % of illicit drug cases involve multiple substances [5]. In Europe, NPS products may contain mixtures of up to seven compounds, with synthetic

* Corresponding author at: Department of Emergency Medicine, National Taiwan University Hospital, 7 Chung Shan S. Rd., Zhongzheng Dist., Taipei City 100225, Taiwan.
E-mail address: conrad@ntu.edu.tw (C.C. Fang).
[1] These authors contributed equally to this study.
[2] TEDAS Research Group also includes the following people: Fang-Cherng Chang, Yen-Chia Chen, Hsien-Yi Chen, Ming-Tai Cheng, Wen-Chih Fann, Sung-Yuan Hu, Chung-Hsun Lee, Sung-Wei Liu, Shu-Yu Shih, Tsung-Hsien Tsai, Cheng-Han Tsai, Ming-Jen Tsai, Te-Hao Wang, I-Jeng Yeh, and Jiun-Hao Yu.

https://doi.org/10.1016/j.ijmedinf.2025.106031
Received 22 January 2025; Received in revised form 11 May 2025; Accepted 28 June 2025
Available online 30 June 2025
1386-5056/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

cathinones detected in over 25 % of these blends [6]. This complexity blurs traditional toxidrome patterns, which emergency physicians rely on to identify drugs like amphetamines or opioids, thus complicating accurate diagnosis in emergency departments (EDs).

Current state-of-the-art (SOTA) approaches for detecting NPS include advanced analytical techniques and computational methods. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) is the gold standard of analytical technique for detecting illicit drugs including NPS [7,8]. It combines liquid chromatography, which separates complex mixtures based on chemical properties, with tandem mass spectrometry, which identifies compounds by their mass-to-charge ratios, enabling precise identification (high sensitivity) of diverse substances in biological samples like urine [8]. Recent advancements, such as high-resolution mass spectrometry (HRMS) and quadrupole time-of-flight (Q-TOF) systems, have further enhanced the detection of unknown NPS compounds [9]. However, these methods are costly, require specialized equipment and expertise, and have turnaround times of hours to days, rendering them impractical for real-time ED diagnostics [10]. Traditional urine drug screen immunoassays, while rapid and cost-effective, suffer from high false positive or false negative rates for NPS due to their structural diversity and limited cross-reactivity [11,12,13]. Efforts to develop NPS-specific immunoassays are underway but remain limited in scope and validation for polysubstance use [14]. Machine learning (ML) has emerged as a SOTA tool for medical diagnostics, with applications in predicting drug toxicity or classifying substance use based on clinical or chemical data [15,16]. However, ML models for NPS detection in ED settings are scarce and often rely on retrospective data, limiting their real-time utility [17].

In Taiwan, NPS use has surged since the early 2000s [18], with the Taiwan Emergency Department Drug Abuse Surveillance (TEDAS) project, established in 2019, revealing that over 50 % of recreational drugs detected in ED cases are NPS, often in polysubstance combinations [7]. While TEDAS employs LC-MS/MS for comprehensive testing, results take several days, underscoring the need for rapid diagnostic tools. Although self-reported surveys from 2014 indicate a stable lifetime drug use prevalence of 1.4 %, the consumption of "club drugs," particularly NPS, has risen significantly [19,20]. The unpredictable and severe effects of NPS place a substantial burden on EDs [21], necessitating tools for early identification and management of illicit drug use.

This study leverages the TEDAS nationwide registry, which systematically collects clinical data and urine samples from suspected ED cases, to develop an ML-based prediction tool for identifying illicit drug use among ED patients in Taiwan. Unlike LC-MS/MS, our tool provides real-time diagnostic support, and unlike immunoassays, it accounts for the complexity of NPS and polysubstance use. By integrating ML algorithms with clinical data, we aim to create a decision support tool that enhances emergency physicians' ability to predict illicit drug use effectively, addressing the limitations of current SOTA approaches.

2. Materials and methods

2.1. Study design, setting, and population

To develop an ML-based prediction tool for identifying illicit drug use among ED patients in Taiwan, this retrospective cohort study leveraged data from the TEDAS database, which is a nationwide de-identified repository that systematically collects clinical manifestations and urine sample analyses from ED patients suspected of illicit drug use. The study included all patients presenting to the EDs of participating hospitals from 2020 to 2023 who underwent urine drug screening due to suspected illicit drug use. Suspicion was based on at least one of 14 standardized clinical criteria, established through expert consensus [7], including bizarre behavior, first-time seizure, abnormal gaze, sympathomimetic toxidrome, altered mental status, out-of-hospital cardiac arrest (OHCA) with no clear cause, suicide, self-harm, violent behavior, jumping from height, motor vehicle accident, self-reported drug use, suspected drug use, or drug use history. The methodological framework consisted of four key steps: (1) defining the study cohort using clinical suspicion criteria, (2) extracting relevant clinical and demographic features from the TEDAS database, (3) preprocessing data for ML model training, and (4) validating model performance against LC-MS/MS-confirmed outcomes. This study was approved by the Institutional Review Board of Chang Gung Memorial Hospital (No. 202300050B0), with a waiver of informed consent due to the retrospective, de-identified nature of the data. The study protocol adhered to the "Minimum Information about Clinical Artificial Intelligence Modeling: MI-CLAIM checklist" guidelines to ensure transparency and reproducibility [22].

2.2. Collection of features

The TEDAS database provided 41 input features, organized into six clinically relevant feature categories—(1) demographics (age, sex), (2) triage data (temperature, heart rate, respiratory rate, systolic and diastolic blood pressure, oxygen saturation, level of consciousness, injury status, transfer, and alcohol use), (3) source of referral (patient, friend, family, emergency medical technician [EMT], police, others), (4) symptoms (seizure, urinary incontinence, facial flushing, sweating, delusion), (5) physical findings (breathing sounds, bowel sounds, myoclonus, pupil response), and (6) the 14 clinical criteria for suspected illicit drug use. These features were chosen for their relevance to ED drug-related presentations.

2.3. Outcome Measures

Urine samples from patients were collected and tested for illicit drug use using electrospray ionization LC-MS/MS [7]. The primary outcome of this study was a positive result for at least one illicit drug detected in the urine sample, which served as the classification label for our binary classification task.

2.4. Data analysis and feature selection

Data entry and processing were performed using Microsoft Excel 2016, with further analysis conducted in IBM SPSS Statistics for Windows (Version 24.0, Armonk, NY, USA). Starting with 41 features, we employed backward sequential feature selection (SFS) to iteratively remove the least impactful features while evaluating metrics such as accuracy, sensitivity, and specificity. This process reduced the feature set to the most predictive variables, enhancing model interpretability and computational efficiency. Features underwent preprocessing, including normalization of continuous variables, encoding of categorical variables, and handling of missing values, to prepare the dataset for ML training. For variables with missing values, we imputed missing data by replacing values with the mean, median, or mode of each respective class, provided the missing rate was below 20 % for that variable. Results are presented as percentages for categorical variables, standard deviations (SD) for continuous variables, and medians with interquartile ranges (IQR) for time variables. Univariate analyses were conducted to assess outcome differences between groups using appropriate statistical tests, including Student's *t*-test, chi-squared test, Fisher's exact test, or Mann–Whitney *U* test, based on data distribution.

2.5. Machine learning model construction

Three supervised ML algorithms—random forest (RF), CatBoost (CB), and light gradient boosting machine (LGBM)—were used to construct prediction models. The dataset was chronologically split into training/validation (2020–2022) and testing cohorts (2023) based on patient presentation time. K-fold cross-validation was applied, with K ranging from 7 to 10, and the model with the highest AUC was selected. For each model, recursive feature elimination with K-fold cross-

Table 1
The characteristics of the study population of the training/validation and testing cohorts.

Features	Total (n = 13615)	Training cohort (n = 9529)	Testing cohort (n = 4086)	P value
Demographics				
Age	38.9 ± 16.3	38.7 ± 16.4	39.3 ± 16.1	0.06
Sex				0.02*
Male	8046 (59.1)	5565 (58.4)	2481 (60.7)	
Female	5453 (40.1)	3874 (40.7)	1579 (38.6)	
NA	116 (0.9)	90 (0.9)	26 (0.6)	
Triage data				
Temperature (℃)	36.7 ± 1.1	36.7 ± 1.2	36.7 ± 0.9	0.03*
Heart rate	99.0 ± 23.4	99.0 ± 23.6	98.9 ± 23.0	0.81
Respiratory rate	19.4 ± 4.7	19.4 ± 4.8	19.4 ± 4.4	0.48
Systolic blood pressure	129.9 ± 26.3	129.9 ± 26.6	130.0 ± 25.6	0.92
Diastolic blood pressure	80.0 ± 18.2	79.9 ± 18.5	80.3 ± 17.7	0.28
Oxygen saturation (%)	97.1 ± 5.5	97.1 ± 5.7	97.2 ± 5.0	0.25
Level of consciousness				<0.001*
Clear	8324 (61.1)	5542 (58.2)	2782 (68.1)	
Drowsy or comatose	3164 (23.2)	2322 (24.4)	842 (20.6)	
Delirium	1239 (9.1)	943 (9.9)	296 (7.2)	
NA	888 (6.5)	722 (7.6)	166 (4.1)	
Injury				<0.001*
No	10,404 (76.4)	7316 (76.8)	3088 (75.6)	
Yes	2702 (19.8)	1793 (18.8)	909 (22.2)	
NA	509 (3.7)	420 (4.4)	89 (2.2)	
Transfer				0.5
No	11,949 (87.8)	8302 (87.1)	3647 (89.3)	
Yes	753 (5.5)	532 (5.6)	221 (5.4)	
NA	913 (6.7)	695 (7.3)	218 (5.3)	
Alcohol use				0.02*
No	10,429 (76.6)	7224 (75.8)	3205 (78.4)	
Yes	2135 (15.7)	1534 (16.1)	601 (14.7)	
NA	1051 (7.7)	771 (8.1)	280 (6.9)	
Source of referral				
Patient				<0.001*
No	6746 (49.5)	4963 (52.1)	1783 (43.6)	
Yes	6792 (49.9)	4514 (47.4)	2278 (55.8)	
NA	77 (0.6)	52 (0.5)	25 (0.6)	
Friend				0.01*
No	12,166 (89.4)	8476 (88.9)	3690 (90.3)	
Yes	1372 (10.1)	1001 (10.5)	371 (9.1)	
NA	77 (0.6)	52 (0.5)	25 (0.6)	
Family				<0.001*
No	7826 (57.5)	5373 (56.4)	2453 (60.0)	
Yes	5712 (42.0)	4104 (43.1)	1608 (39.4)	
NA	77 (0.6)	52 (0.5)	25 (0.6)	
EMT				0.87
No	10,320 (75.8)	7228 (75.9)	3092 (75.7)	
Yes	3218 (23.6)	2249 (23.6)	969 (23.7)	
NA	77 (0.6)	52 (0.5)	25 (0.6)	
Police				<0.001*
No	12,097 (88.9)	8579 (90.0)	3518 (86.1)	
Yes	1441 (10.6)	898 (9.4)	543 (13.3)	
NA	77 (0.6)	52 (0.5)	25 (0.6)	
Others				0.51
No	12,877 (94.6)	9003 (94.5)	3874 (94.8)	
Yes	648 (4.8)	461 (4.8)	187 (4.6)	
NA	90 (0.7)	65 (0.7)	25 (0.6)	
Symptoms				
Seizure				<0.001*
No	11,870 (87.2)	8279 (86.9)	3591 (87.9)	
Yes	599 (4.4)	458 (4.8)	141 (3.5)	
NA	1146 (8.4)	792 (8.3)	354 (8.7)	
Urine incontinence				0.03*
No	11,542 (84.8)	7995 (83.9)	3547 (86.8)	
Yes	831 (6.1)	605 (6.3)	226 (5.5)	
NA	1242 (9.1)	929 (9.7)	313 (7.7)	
Facial flushing				<0.001*
No	10,770 (79.1)	7654 (80.3)	3116 (76.3)	
Yes	1849 (13.6)	1189 (12.5)	660 (16.2)	
NA	996 (7.3)	686 (7.2)	310 (7.6)	
Sweating				0.02*
No	11,472 (84.3)	8014 (84.1)	3458 (84.6)	
Yes	1641 (12.1)	1099 (11.5)	542 (13.3)	
NA	502 (3.7)	416 (4.4)	86 (2.1)	
Delusion				<0.001*
No	10,403 (76.4)	7392 (77.6)	3011 (73.7)	
Yes	2011 (14.8)	1285 (13.5)	726 (17.8)	
NA	1201 (8.8)	852 (8.9)	349 (8.5)	
Physical findings				
Breathing sound				<0.001*
Normal	12,612 (92.6)	8853 (92.9)	3759 (92.0)	
Abnormal	634 (4.7)	419 (4.4)	215 (5.3)	
Absent	112 (0.8)	62 (0.7)	50 (1.2)	
NA	257 (1.9)	195 (2.0)	62 (1.5)	
Bowel sound				<0.001*
Normal	11,556 (84.9)	7951 (83.4)	3605 (88.2)	
Slow	691 (5.1)	540 (5.7)	151 (3.7)	
Rapid	387 (2.8)	291 (3.1)	96 (2.3)	
NA	981 (7.2)	747 (7.8)	234 (5.7)	
Myoclonus				<0.001*
No	10,965 (80.5)	7782 (81.7)	3183 (77.9)	

(*continued on next page*)

Table 1 (*continued*)

Features	Total (n = 13615)	Training cohort (n = 9529)	Testing cohort (n = 4086)	P value
Yes	1060 (7.8)	564 (5.9)	496 (12.1)	
NA	1590 (11.7)	1183 (12.4)	407 (10.0)	
Pupils				<0.001*
Normal	11,478 (84.3)	7934 (83.3)	3544 (86.7)	
Mydriasis	504 (3.7)	343 (3.6)	161 (3.9)	
Miosis	692 (5.1)	533 (5.6)	159 (3.9)	
Anisocoric	80 (0.6)	54 (0.6)	26 (0.6)	
NA	861 (6.3)	665 (7.0)	196 (4.8)	
Clinical characteristics suspected of illicit drug use				
Bizarre behavior				0.72
No	9905 (72.8)	6916 (72.6)	2989 (73.2)	
Yes	3551 (26.1)	2468 (25.9)	1083 (26.5)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
First-time seizure				<0.001*
No	13,002 (95.5)	9034 (94.8)	3968 (97.1)	
Yes	454 (3.3)	350 (3.7)	104 (2.5)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
Abnormal gaze pattern				0.46
No	10,910 (80.1)	7593 (79.7)	3317 (81.2)	
Yes	2546 (18.7)	1791 (18.8)	755 (18.5)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
Sympathomimetic toxidrome				<0.001*
No	12,720 (93.4)	9029 (94.8)	3691 (90.3)	
Yes	736 (5.4)	355 (3.7)	381 (9.3)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
Altered mental status				<0.001*
No	10,167 (74.7)	6903 (72.4)	3264 (79.9)	
Yes	3289 (24.2)	2481 (26.0)	808 (19.8)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
OHCA without clue				0.007*
No	13,302 (97.7)	9292 (97.5)	4010 (98.1)	
Yes	154 (1.1)	92 (1.0)	62 (1.5)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
Suicide				0.01*
No	11,359 (83.4)	7872 (82.6)	3487 (85.3)	
Yes	2097 (15.4)	1512 (15.9)	585 (14.3)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
Self-harm				0.05
No	12,289 (90.3)	8599 (90.2)	3690 (90.3)	
Yes	1167 (8.6)	785 (8.2)	382 (9.3)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
Violent behavior				<0.001*
No	10,827 (79.5)	7290 (76.5)	3537 (86.6)	
Yes	2699 (19.8)	2164 (22.7)	535 (13.1)	
NA	89 (0.7)	75 (0.8)	14 (0.3)	
Jumping from height				0.24
No	13,306 (97.7)	9286 (97.4)	4020 (98.4)	
Yes	150 (1.1)	98 (1.0)	52 (1.3)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
Motor vehicle accident				<0.001*
No	12,717 (93.4)	8930 (93.7)	3787 (92.7)	
Yes	739 (5.4)	454 (4.8)	285 (7.0)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
Self-reported drug use				<0.001*
No	12,166 (89.4)	8350 (87.6)	3816 (93.4)	
Yes	1290 (9.5)	1034 (10.9)	256 (6.3)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
Suspicion of drug use				<0.001*
No	8965 (65.8)	6530 (68.5)	2435 (59.6)	
Yes	4491 (33.0)	2854 (30.0)	1637 (40.1)	
NA	159 (1.2)	145 (1.5)	14 (0.3)	
History of drug use				<0.001*
No	4236 (31.1)	2811 (29.5)	1425 (34.9)	
Yes	8916 (65.5)	6375 (66.9)	2541 (62.2)	
NA	463 (3.4)	343 (3.6)	120 (2.9)	

Abbreviations: OHCA, out-of-hospital cardiac arrest; EMT, emergency medical technician; NA, not available.
Note: * $P < 0.05$.

validation within the training set identified key features, iteratively removing features until AUC declined. The features selected by the top-performing model were then applied to the other models, and performance was evaluated.

Alongside AUC, we reported classification performance on the testing cohort, including accuracy, F1-score, precision (positive predictive value [PPV]), recall (sensitivity), specificity, negative predictive value (NPV), and area under the precision-recall curve (AUPRC) for each model. To enhance interpretability, SHapley Additive exPlanations (SHAP) analysis and feature importance were incorporated [23]. All ML analyses were conducted in Python 3.8 (Python Software Foundation, Wilmington, DE) with the scikit-learn 0.23.1 package. Additionally, a web-based decision support tool and mobile apps were developed using the ML model, providing free nationwide access for emergency physicians to early identify illicit drug use in emergency settings.

3. Results

3.1. Characteristics of study subjects

From 2020 to 2023, the TEDAS project recorded 13,615 ED visits with suspected drug use, identifying 3,185 positive cases (23.4 %). The dataset was split into a training/validation cohort (9,529 cases, 2020–2022; 2,371 positive, 24.9 %) and a testing cohort (4,086 cases, 2023; 814 positive, 19.9 %). Table 1 summarizes the demographic and clinical characteristics of both cohorts, while Table 2 presents the univariate analyses of features for patients with and without illicit drug use.

3.2. Main results

Using 20 features selected via backward SFS and 7-fold cross-validation, three ML models—RF, CB, and LGBM—were developed to predict illicit drug use. the LGBM model initially achieved the best AUC performance. These same features were then applied to the other two models (CB and RF). The selected 20 features included demographics (age, sex); triage data (temperature, heart rate, respiratory rate, systolic

Table 2
Characteristics and univariate analyses of features between patients with or without illicit drugs use on the training and testing cohorts.

Features	Training cohort (n = 9529)		P value	Testing cohort (n = 4086)		P value
	Urine test (−) (n = 7158)	Urine test (+) (n = 2371)		Urine test (−) (n = 3272)	Urine test (+) (n = 814)	
Demographics						
Age	39.6 ± 17.4	36.0 ± 12.2	<0.001*	39.6 ± 17.0	38.0 ± 12.1	0.01*
Sex			<0.001*			<0.001*
Male	3882 (54.2)	1683 (71.0)		1873 (57.2)	608 (74.7)	
Female	3219 (45.0)	655 (27.6)		1377 (42.1)	202 (24.8)	
NA	57 (0.8)	33 (1.4)		22 (0.7)	4 (0.5)	
Triage data						
Temperature (℃)	36.7 ± 1.2	36.8 ± 1.2	0.001*	36.7 ± 0.9	36.7 ± 1.1	0.72
Heart rate	97.2 ± 22.7	104.6 ± 25.6	<0.001*	97.5 ± 22.6	104.6 ± 23.5	<0.001*
Respiratory rate	19.4 ± 4.8	19.6 ± 4.6	0.1	19.4 ± 4.5	19.4 ± 3.9	0.69
Systolic blood pressure	129.8 ± 26.6	130.4 ± 26.4	0.36	129.3 ± 25.4	132.4 ± 26.4	0.003*
Diastolic blood pressure	79.5 ± 18.0	81.4 ± 19.7	<0.001*	79.4 ± 17.1	84.1 ± 19.4	<0.001*
Oxygen saturation (%)	97.1 ± 6.0	97.1 ± 4.4	0.95	97.4 ± 4.4	96.3 ± 6.7	<0.001*
Level of consciousness			<0.001*			<0.001*
Clear	4167 (58.2)	1375 (58.0)		2297 (70.2)	485 (59.6)	
Drowsy or comatose	1816 (25.4)	506 (21.3)		653 (20.0)	189 (23.2)	
Delirium	657 (9.2)	286 (12.1)		206 (6.3)	90 (11.1)	
NA	518 (7.2)	204 (8.6)		116 (3.5)	50 (6.1)	
Injury			0.001*			0.05
No	5578 (77.9)	1738 (73.3)		2456 (75.1)	632 (77.6)	
Yes	1302 (18.2)	491 (20.7)		750 (22.9)	159 (19.5)	
NA	278 (3.9)	142 (6.0)		66 (2.0)	23 (2.8)	
Transfer			0.02*			0.79
No	6313 (88.2)	1989 (83.9)		2931 (89.6)	716 (88.0)	
Yes	380 (5.3)	152 (6.4)		176 (5.4)	45 (5.5)	
NA	465 (6.5)	230 (9.7)		165 (5.0)	53 (6.5)	
Alcohol use			0.04*			0.12
No	5458 (76.3)	1766 (74.5)		2569 (78.5)	636 (78.1)	
Yes	1196 (16.7)	338 (14.3)		498 (15.2)	103 (12.7)	
NA	504 (7.0)	267 (11.3)		205 (6.3)	75 (9.2)	
Source of referral						
Patient			0.5			0.26
No	3743 (52.3)	1220 (51.5)		1414 (43.2)	369 (45.3)	
Yes	3377 (47.2)	1137 (48.0)		1839 (56.2)	439 (53.9)	
NA	38 (0.5)	14 (0.6)		19 (0.6)	6 (0.7)	
Friend			0.09			0.21
No	6390 (89.3)	2086 (88.0)		2965 (90.6)	725 (89.1)	
Yes	730 (10.2)	271 (11.4)		288 (8.8)	83 (10.2)	
NA	38 (0.5)	14 (0.6)		19 (0.6)	6 (0.7)	
Family			<0.001*			<0.001*
No	3897 (54.4)	1476 (62.3)		1913 (58.5)	540 (66.3)	
Yes	3223 (45.0)	881 (37.2)		1340 (41.0)	268 (32.9)	
NA	38 (0.5)	14 (0.6)		19 (0.6)	6 (0.7)	
EMT			<0.001*			0.11
No	5498 (76.8)	1730 (73.0)		2494 (76.2)	598 (73.5)	
Yes	1622 (22.7)	627 (26.4)		759 (23.2)	210 (25.8)	
NA	38 (0.5)	14 (0.6)		19 (0.6)	6 (0.7)	
Police			<0.001*			<0.001*
No	6586 (92.0)	1993 (84.1)		2861 (87.4)	657 (80.7)	
Yes	534 (7.5)	364 (15.4)		392 (12.0)	151 (18.6)	
NA	38 (0.5)	14 (0.6)		19 (0.6)	6 (0.7)	
Others			0.04*			0.004*
No	6749 (94.3)	2254 (95.1)		3088 (94.4)	786 (96.6)	
Yes	365 (5.1)	96 (4.0)		165 (5.0)	22 (2.7)	
NA	44 (0.6)	21 (0.9)		19 (0.6)	6 (0.7)	
Symptoms						
Seizure			0.02*			0.89
No	6230 (87.0)	2049 (86.4)		2887 (88.2)	704 (86.5)	
Yes	366 (5.1)	92 (3.9)		114 (3.5)	27 (3.3)	
NA	562 (7.9)	230 (9.7)		271 (8.3)	83 (10.2)	
Urine incontinence			0.51			0.24
No	6036 (84.3)	1959 (82.6)		2859 (87.4)	688 (84.5)	
Yes	464 (6.5)	141 (5.9)		175 (5.3)	51 (6.3)	
NA	658 (9.2)	271 (11.4)		238 (7.3)	75 (9.2)	
Facial flushing			0.009*			0.58
No	5824 (81.4)	1830 (77.2)		2508 (76.7)	608 (74.7)	

(*continued on next page*)

Table 2 (*continued*)

Features	Training cohort (n = 9529)		P value	Testing cohort (n = 4086)		P value
	Urine test (−) (n = 7158)	Urine test (+) (n = 2371)		Urine test (−) (n = 3272)	Urine test (+) (n = 814)	
Yes	863 (12.1)	326 (13.7)		525 (16.0)	135 (16.6)	
NA	471 (6.6)	215 (9.1)		239 (7.3)	71 (8.7)	
Sweating			<0.001*			<0.001*
No	6181 (86.4)	1833 (77.3)		2818 (86.1)	640 (78.6)	
Yes	688 (9.6)	411 (17.3)		395 (12.1)	147 (18.1)	
NA	289 (4.0)	127 (5.4)		59 (1.8)	27 (3.3)	
Delusion			<0.001*			0.01*
No	5677 (79.3)	1715 (72.3)		2443 (74.7)	568 (69.8)	
Yes	863 (12.1)	422 (17.8)		560 (17.1)	166 (20.4)	
NA	618 (8.6)	234 (9.9)		269 (8.2)	80 (9.8)	
Physical findings						
Breathing sound			0.01*			0.004*
Normal	6670 (93.2)	2183 (92.1)		3033 (92.7)	726 (89.2)	
Abnormal	302 (4.2)	117 (4.9)		154 (4.7)	61 (7.5)	
Absent	38 (0.5)	24 (1.0)		38 (1.2)	12 (1.5)	
NA	148 (2.1)	47 (2.0)		47 (1.4)	15 (1.8)	
Bowel sound			0.003*			0.01*
Normal	6031 (84.3)	1920 (81.0)		2907 (88.8)	698 (85.7)	
Slow	410 (5.7)	130 (5.5)		126 (3.9)	25 (3.1)	
Rapid	195 (2.7)	96 (4.0)		66 (2.0)	30 (3.7)	
NA	522 (7.3)	225 (9.5)		173 (5.3)	61 (7.5)	
Myoclonus			<0.001*			0.66
No	5965 (83.3)	1817 (76.6)		2562 (78.3)	621 (76.3)	
Yes	365 (5.1)	199 (8.4)		395 (12.1)	101 (12.4)	
NA	828 (11.6)	355 (15.0)		315 (9.6)	92 (11.3)	
Pupils			<0.001*			<0.001*
Normal	6128 (85.6)	1806 (76.2)		2891 (88.4)	653 (80.2)	
Mydriasis	201 (2.8)	142 (6.0)		118 (3.6)	43 (5.3)	
Miosis	304 (4.2)	229 (9.7)		91 (2.8)	68 (8.4)	
Anisocoric	39 (0.5)	15 (0.6)		24 (0.7)	2 (0.2)	
NA	486 (6.8)	179 (7.5)		148 (4.5)	48 (5.9)	
Clinical characteristics suspected of illicit drug use						
Bizarre behavior			0.03*			0.88
No	5241 (73.2)	1675 (70.6)		2392 (73.1)	597 (73.3)	
Yes	1815 (25.4)	653 (27.5)		869 (26.6)	214 (26.3)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
First-time seizure			0.01*			0.75
No	6773 (94.6)	2261 (95.4)		3179 (97.2)	789 (96.9)	
Yes	283 (4.0)	67 (2.8)		82 (2.5)	22 (2.7)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
Abnormal gaze pattern			0.05			0.64
No	5741 (80.2)	1852 (78.1)		2661 (81.3)	656 (80.6)	
Yes	1315 (18.4)	476 (20.1)		600 (18.3)	155 (19.0)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
Sympathomimetic toxidrome			<0.001*			<0.001*
No	6850 (95.7)	2179 (91.9)		2997 (91.6)	694 (85.3)	
Yes	206 (2.9)	149 (6.3)		264 (8.1)	117 (14.4)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
Altered mental status			0.09			0.02*
No	5159 (72.1)	1744 (73.6)		2637 (80.6)	627 (77.0)	
Yes	1897 (26.5)	584 (24.6)		624 (19.1)	184 (22.6)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
OHCA without clue			0.21			0.03*
No	6992 (97.7)	2300 (97.0)		3218 (98.3)	792 (97.3)	
Yes	64 (0.9)	28 (1.2)		43 (1.3)	19 (2.3)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
Suicide			<0.001*			<0.001*
No	5828 (81.4)	2044 (86.2)		2746 (83.9)	741 (91.0)	
Yes	1228 (17.2)	284 (12.0)		515 (15.7)	70 (8.6)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
Self-harm			0.59			<0.001*
No	6472 (90.4)	2127 (89.7)		2929 (89.5)	761 (93.5)	
Yes	584 (8.2)	201 (8.5)		332 (10.1)	50 (6.1)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
Violent behavior			<0.001*			0.45
No	5544 (77.5)	1746 (73.6)		2826 (86.4)	711 (87.3)	
Yes	1560 (21.8)	604 (25.5)		435 (13.3)	100 (12.3)	
NA	54 (0.8)	21 (0.9)		11 (0.3)	3 (0.4)	
Jumping from height			0.01*			0.64

(*continued on next page*)

Table 2 (*continued*)

Features	Training cohort (n = 9529)		P value	Testing cohort (n = 4086)		P value
	Urine test (−) (n = 7158)	Urine test (+) (n = 2371)		Urine test (−) (n = 3272)	Urine test (+) (n = 814)	
No	6993 (97.7)	2293 (96.7)		3218 (98.3)	802 (98.5)	
Yes	63 (0.9)	35 (1.5)		43 (1.3)	9 (1.1)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
Motor vehicle accident			0.08			0.006*
No	6699 (93.6)	2231 (94.1)		3015 (92.1)	772 (94.8)	
Yes	357 (5.0)	97 (4.1)		246 (7.5)	39 (4.8)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
Self-reported drug use			<0.001*			<0.001*
No	6450 (90.1)	1900 (80.1)		3153 (96.4)	663 (81.4)	
Yes	606 (8.5)	428 (18.1)		108 (3.3)	148 (18.2)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
Suspicion of drug use			<0.001*			0.2
No	5051 (70.6)	1479 (62.4)		1966 (60.1)	469 (57.6)	
Yes	2005 (28.0)	849 (35.8)		1295 (39.6)	342 (42.0)	
NA	102 (1.4)	43 (1.8)		11 (0.3)	3 (0.4)	
History of drug use			<0.001*			<0.001*
No	2329 (32.5)	482 (20.3)		1250 (38.2)	175 (21.5)	
Yes	4573 (63.9)	1802 (76.0)		1923 (58.8)	618 (75.9)	
NA	256 (3.6)	87 (3.7)		99 (3.0)	21 (2.6)	

Abbreviations: OHCA, out-of-hospital cardiac arrest; EMT, emergency medical technician; NA, not available.
Note: * $P < 0.05$.

and diastolic blood pressure, oxygen saturation, level of consciousness [clear, delirium, drowsy, comatose], transfer status, and alcohol use); source of referral (patient, EMT); physical findings (pupil status [normal, miosis, mydriasis, anisocoria]); and clinical characteristics (bizarre behavior, sympathomimetic toxidrome, suicide, self-reported drug use, suspected drug use, and history of drug use). The models demonstrated strong discrimination in the training/validation cohort and maintained robust AUC performance in the testing cohort (Fig. 1).

The models were trained and validated using 7-fold cross-validation on the training/validation cohort (9,529 cases, 2020–2022) and evaluated on a separate testing cohort (4,086 cases, 2023), with illicit drug use outcomes confirmed by LC-MS/MS, the gold standard for drug detection. Of them, the CB model achieved the highest AUC of 0.846 (95 % CI: 0.831–0.859) in the testing cohort, followed by LGBM at 0.843 (95 % CI: 0.829–0.857) and RF at 0.828 (95 % CI: 0.813–0.843). For precision-recall balance, CB and LGBM both scored an AUPRC of 0.663 (95 % CI: 0.634–0.692), with RF at 0.622 (95 % CI: 0.591–0.652). No significant differences in AUC or AUPRC were observed across the models. Additional metrics (accuracy, sensitivity, specificity, NPV, and PPV) are detailed in Table 3. Hyperparameters tuned for each model are detailed in Supplementary Table 1 (20 features) and Supplementary Table 2 (41 features).

Fig. 2 presents a heat map of the computed 20 features, ordered by their median normalized importance across all three constructed ML models. Fig. 3 illustrates feature importance using SHAP values, with outcomes analyzed through 7-fold cross-validation, visually highlighting the impact of each feature on prediction accuracy within the models. A web-based decision support tool (https://www.toxconsult.org.tw/calculator, in both Chinese and English) and mobile apps (Android: https://reurl.cc/x11OxE; iOS: https://reurl.cc/nZZEVv, in Chinese) were developed to provide free access to the prediction algorithms.

4. Discussion

Our ML models—RF, CB, and LGBM—demonstrated robust performance in predicting illicit drug use among ED patients, with CB achieving the highest AUC of 0.846 (95 % CI: 0.831–0.859) and a PPV of 78.5 % in the testing cohort. This represents a significant improvement over clinical suspicion, highlighting the models' ability to enhance diagnostic accuracy. The use of LC-MS/MS as the ground truth and temporal cohort separation confirms the models' effectiveness for ED diagnostics. By integrating routinely collected ED clinical data from patient history and physical exams, our models enhance their applicability for real-time diagnostics and enable physicians to initiate treatments.

Instead of using a traditional ML model, all three models were designed with ensemble methods, which combine multiple predictive models to create a single, unified prediction, thereby enhancing accuracy [24]. RF was chosen for its robustness and interpretability, capturing non-linear feature interactions in the TEDAS data and reducing overfitting [25]. CB was selected for its superior handling of categorical features minimizing preprocessing and overfitting through ordered boosting [26], and LGBM was included for its computational efficiency and scalability, enabling rapid predictions critical for ED use [27]. Our models showed excellent discriminatory performance, particularly in terms of AUC on the testing cohort. All model achieved high specificity and NPV, supporting their utility in reliably identifying ED patients unlikely to have used illicit drug use —potentially reducing the need for costly confirmatory laboratory tests. The PPV and sensitivity were also satisfactory, indicating that the models can both rule out non-drug-using patients and assist in guiding early clinical decision-making for suspected users. Among4,086 cases in the testing cohort (814 positives for drug), the pre-test probability of drug use based on physician judgment was approximately 20 %. In comparison, the CB classifier, the best-performing model, achieved an AUC of 0.887 (95 % CI: 0.879–0.894) in the training set and 0.846 (95 % CI: 0.831–0.859) in the testing set, with a PPV of 78.5 %. This represents a marked improvement over baseline clinical estimates, reinforcing the model's potential to enhance early risk stratification and clinical decision-making in the ED setting [28,29].

The rising prevalence of polysubstance use—evident in Taiwan (e.g., cathinones with ketamine, methamphetamine, opioids) and globally (e.g., opioid-methamphetamine combinations in the U.S.)—complicates diagnosis by obscuring traditional toxidrome patterns [30–36]. Our models, trained on TEDAS data validated by LC-MS/MS, effectively predict complex drug use patterns, offering a robust tool for EDs where polysubstance use is increasingly common. In our ML model, demographic factors like age and sex are key features, while vital signs—including GCS and oxygen saturation—play a critical role, especially during ED triage. Patient history elements, such as self-reported illicit drug use, also strongly predict drug use, whereas combined alcohol use

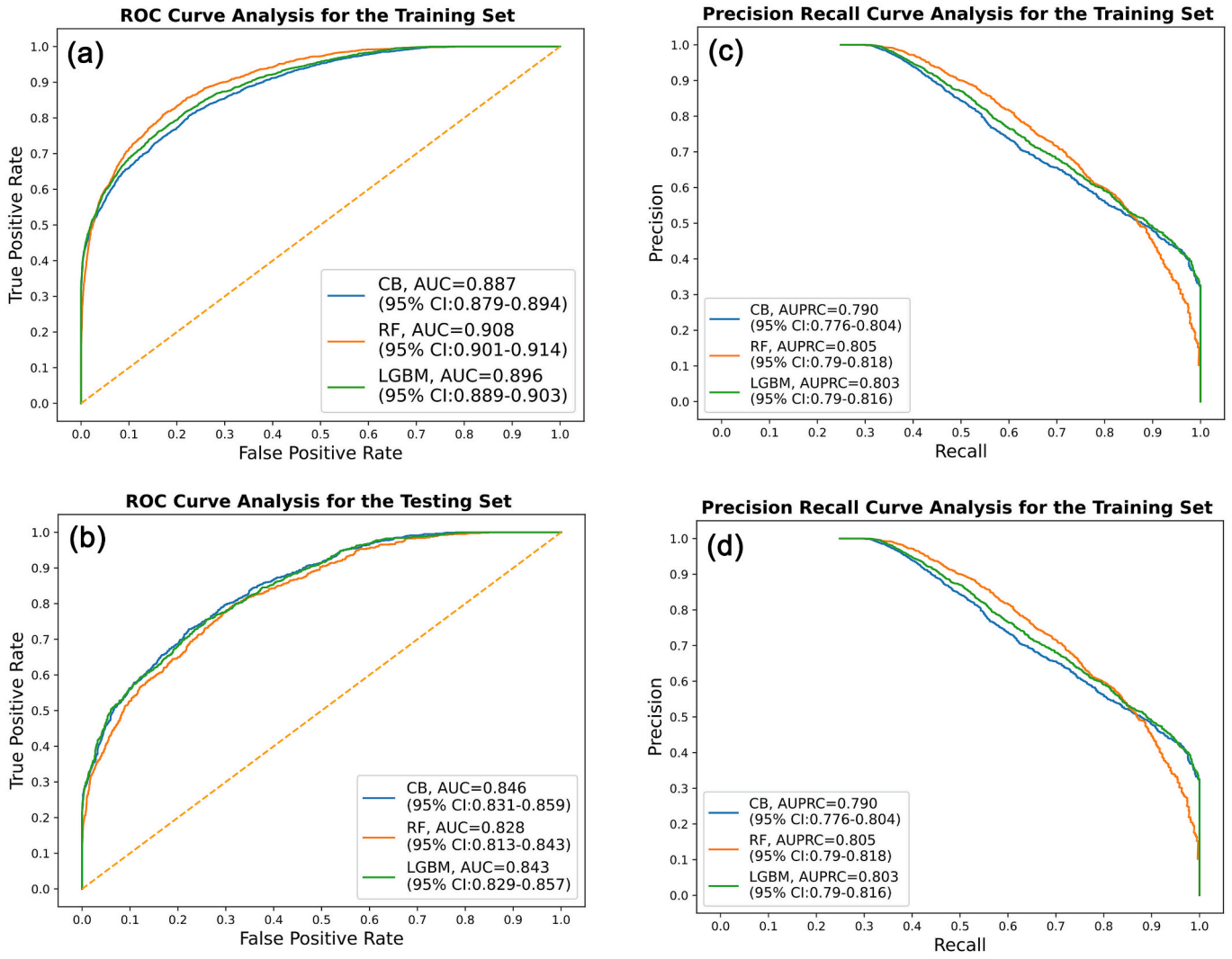


Fig. 1. Results and the comparisons of the three ML models based on the performances of area under the receiver operating characteristic curves (AUC) on the Training/Validation (A), and Testing (B) cohorts; and area under the precision-recall curve (AUPRC) on the Training/Validation (C), and Testing (D) cohorts. (Abbreviations: RF, random forest; CB, CatBoost; LGBM, light gradient boosting machine).

Table 3
Comparison between model performances on the training/validation and testing cohorts under 7-fold cross validation.

Model	AUC (95 % CI)	AUPRC (95 % CI)	Accuracy	F1	Kappa	Sensitivity (Recall)	Specificity	PPV (Precision)	NPV
Training/Validation cohort									
CatBoost	0.887 (0.879–0.894)	0.79 (0.776–0.804)	0.855	0.612	0.534	0.460	0.986	0.914	0.846
LGBT	0.896 (0.889–0.903)	0.803 (0.79–0.816)	0.858	0.626	0.549	0.477	0.985	0.913	0.85
Random forest	0.908 (0.901–0.914)	0.805 (0.79–0.818)	0.834	0.700	0.587	0.776	0.854	0.637	0.920
Testing cohort									
CatBoost	0.846 (0.831–0.859)	0.663 (0.634–0.692)	0.852	0.489	0.417	0.355	0.976	0.785	0.859
LGBT	0.843 (0.829–0.857)	0.664 (0.635–0.693)	0.856	0.519	0.445	0.391	0.972	0.774	0.865
Random forest	0.828 (0.813–0.843)	0.622 (0.591–0.652)	0.777	0.535	0.394	0.645	0.810	0.457	0.902

Abbreviations: LGBT, light gradient boosting machine; AUC, area under the receiver operating characteristic curve; AUPRC, area under the precision recall curve; PPV, positive predictive value; NPV, negative predictive value.

and presentations for suicide act as negative indicators. Among clinical signs, miosis, mydriasis, delirium, and sympathetic toxidromes are strong indicators of illicit drug use. A notable advantage of our ML models is their reliance on routinely collected clinical data, making them highly applicable for ED physicians to support initial diagnosis and treatment decisions.

Compared to the current SOTA methods, our ML models provide rapid, cost-effective predictions of illicit drug use, generating results in seconds by processing routine clinical data via a web-based tool or mobile app, making them ideal for real-time ED diagnostics. In contrast, LC-MS/MS, the gold standard for detecting NPS, requires hours to days for sample preparation and analysis, limiting its practicality due to high costs and specialized equipment needs [8]. Urine drug screen immunoassays, another SOTA method, offer results in 30–60 min but suffer

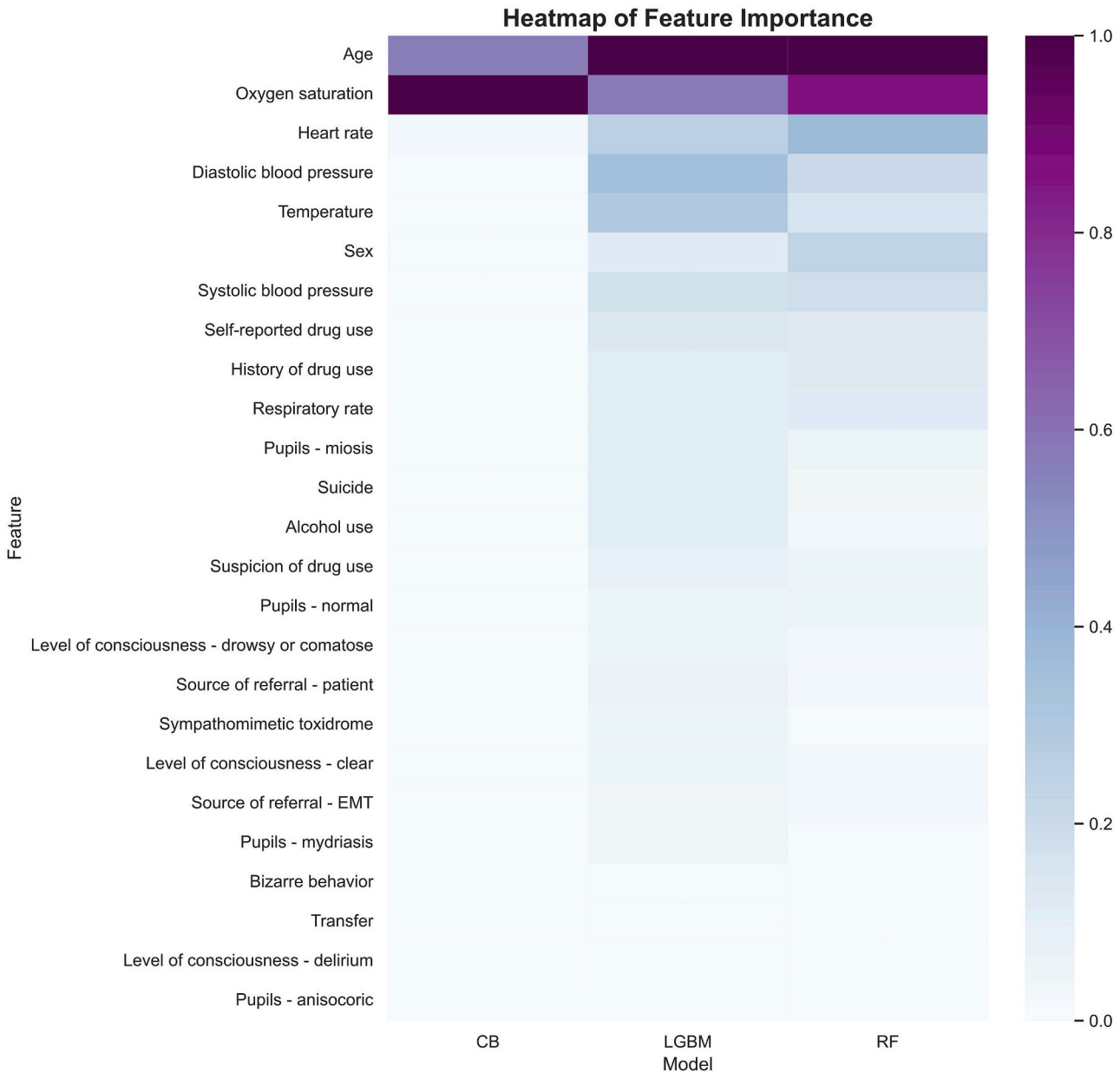


Fig. 2. A heat map of the selected 20 features ordered by computed median normalized importance across all models of the constructed five different ML models.

from inaccuracy for NPS, reducing their reliability [11–13]. Our models, leveraging routine clinical data, enable rapid, accurate predictions, addressing the limitations of these methods and supporting ED decision-making in the NPS era.

5. Limitations

This study faces several limitations. First, there is a potential for selection bias, as the analysis included only TEDAS cases with urine drug samples meeting specific criteria, potentially missing some illicit drug users and underestimate prevalence, which may limit model applicability to all ED cases. Second, manual data entry into the TEDAS registry risks errors, incomplete records, or inconsistencies potentially affecting model's accuracy and generalizability. Third, the Taiwan-centric dataset may limit the findings' applicability to regions with different drug use patterns, healthcare systems, or demographics. Future validation across diverse populations is needed.

6. Conclusions

This study demonstrates the efficacy of ML model in predicting illicit drug use among ED patients, achieving a robust AUC of 0.846 and a PPV of 78.5 % with CB classifier, significantly outperforming clinical suspicion. By leveraging routine clinical data, these models address the diagnostic challenges posed by NPS and polysubstance use, which obscure traditional toxidrome patterns. The models' high specificity and negative predictive value enable reliable identification of non-drug-using patients, reducing unnecessary laboratory tests, while their sensitivity supports early intervention for drug-related conditions. Freely accessible via a web-based tool and mobile apps, the models offer a cost-effective, real-time diagnostic solution for ED physicians worldwide. Their potential to lower healthcare costs and improve patient outcomes is particularly significant in resource-constrained settings. However, further validation across diverse international populations and prospective studies are essential to ensure adaptability to varying drug use patterns and healthcare systems. These steps will solidify the

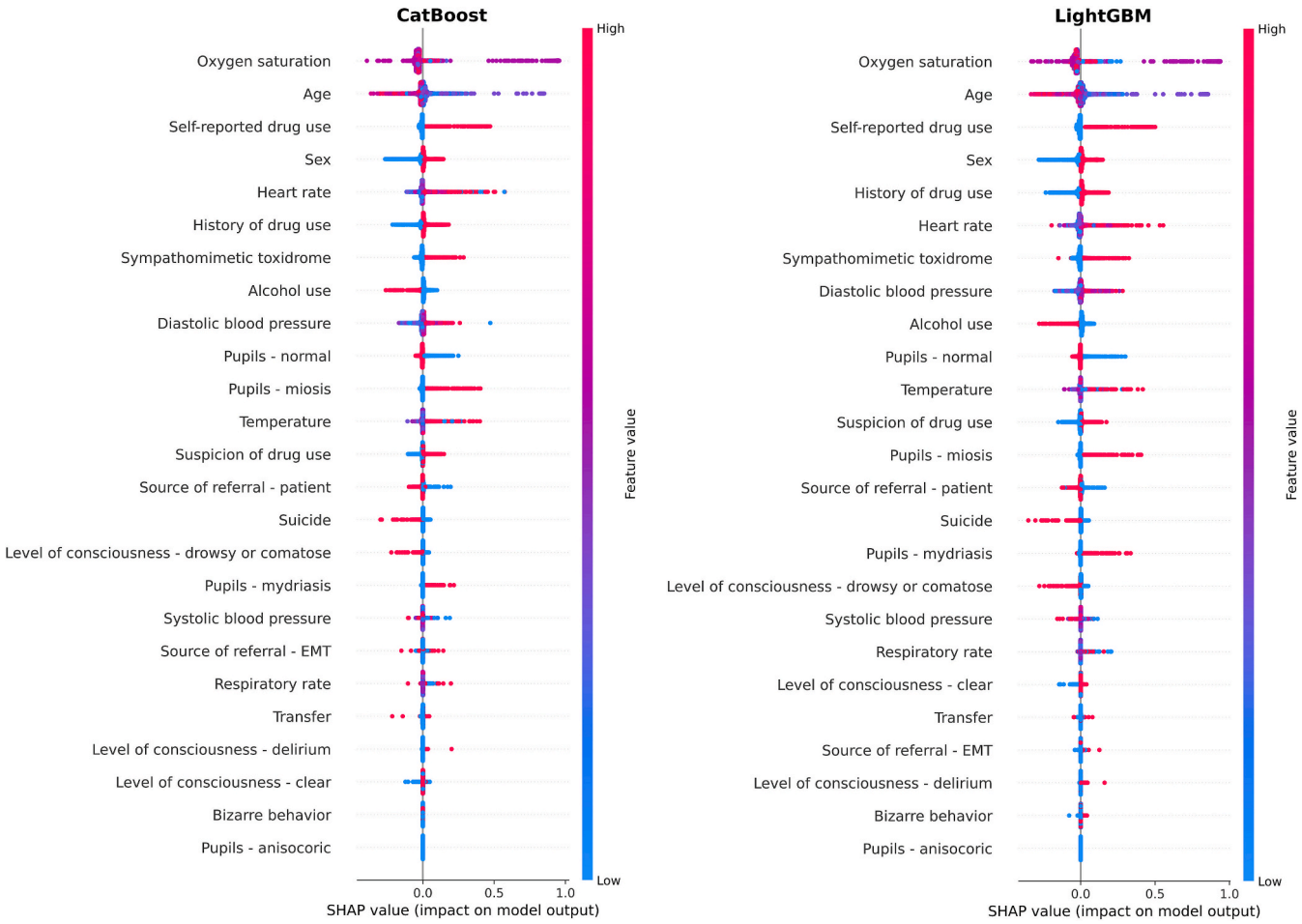


Fig. 3. The SHapley Additive exPlanations (SHAP) values of the 20 important features as a way to explain the output of the constructed machine learning models by selecting 7-fold cross-validation using CB classifiers (A), LGBM (B), and RF (C). (Abbreviations: RF, random forest; CB, CatBoost; LGBM, light gradient boosting machine).

models' role in transforming ED diagnostics in the era of complex illicit drug use.

Summary points

What was already known on the topic?

- Identifying illicit drug use through urine testing in emergency departments (EDs) is time-consuming and often challenging, particularly with the emergence of new psychoactive substances.
- Conventional methods for detecting illicit drug use rely heavily on laboratory testing, which may delay clinical decision-making.
- Machine learning (ML) models have been explored for various clinical applications, but their utility in predicting illicit drug use in ED patients has not been extensively studied.

What this study added to our knowledge?

- This study developed a machine learning prediction model using data from the Taiwan Emergency Department Drug Abuse Surveillance (TEDAS) database, incorporating six categories of clinical features for early identification of illicit drug use.
- The CatBoost classifier demonstrated robust predictive performance with an AUC of 0.846, outperforming other ML algorithms in the testing cohort.
- A practical application of the ML model was achieved through the development of a web-based tool and mobile apps, offering free and accessible decision support for emergency physicians.
- The study provides a foundation for broader implementation and further validation in international settings, highlighting the potential for ML to streamline and enhance clinical workflows in detecting drug abuse.

Author statement

This study was approved by the Institutional Review Board of Chang Gung Memorial Hospital (No. 202300050B0) with a waiver of informed consent since this is a retrospective study without any identifiable patient information. The authors also declare that the work described has not been published previously (except in the form of an abstract), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copy right holder.

Funding

This work was supported by a grant from the Taiwan Food and Drug Administration (grant 108TFDA-N-003).

CRediT authorship contribution statement

Tsung-Chien Lu: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Chih-Chuan Lin:** Writing – original draft, Methodology, Data curation,

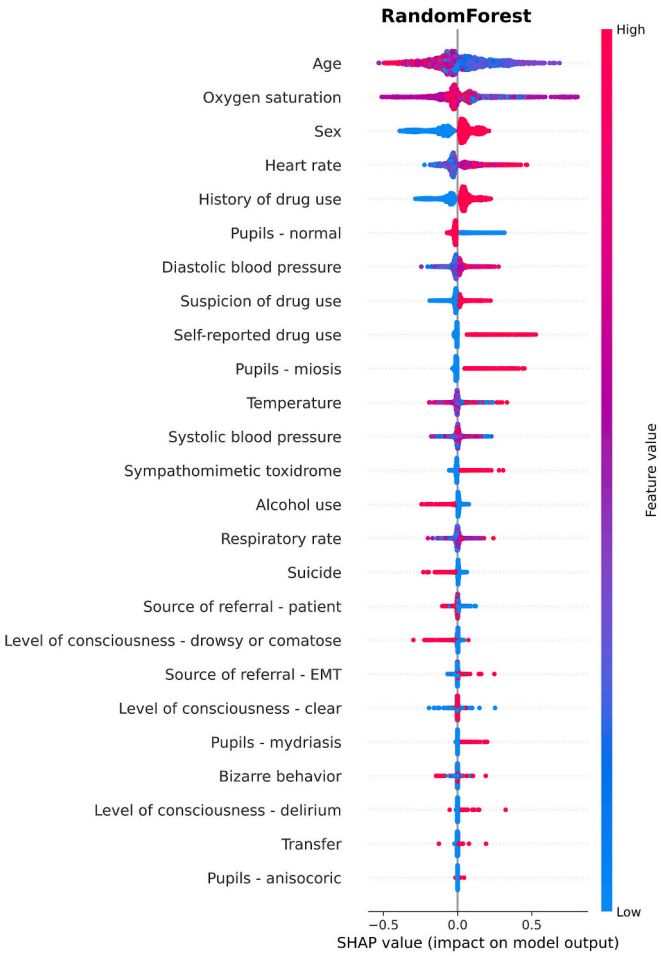


Fig. 3. (*continued*).

Conceptualization. **Te-I Weng**: Writing – review & editing, Data curation. **Fan-Ya Chou:** Validation, Software, Methodology, Investigation, Formal analysis. **Cheng-Chung Fang:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

None.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2025.106031.

References

[1] European Monitoring Centre for Drugs and Drug Addiction. European Drug Report 2023: Trends and Developments. Accessed October 16, 2024. https://www.emcdda.europa.eu/publications/european-drug-report/2023_en.
[2] United Nations Office on Drugs and Crime. World drug report 2021. Accessed October 16, 2024. https://www.unodc.org/unodc/en/data-and-analysis/wdr2021.html.
[3] J. Lee, S. Yang, Y. Kang, E. Han, L.Y. Feng, J.H. Li, et al., Prevalence of new psychoactive substances in Northeast Asia from 2007 to 2015, Forensic Sci. Int. 272 (2017) 1–9, https://doi.org/10.1016/j.forsciint.2016.10.010.
[4] H. Chung, J. Lee, E. Kim, Trends of novel psychoactive substances (NPSs) and their fatal cases, Forensic Toxicol. 34 (2016) 1–11, https://doi.org/10.1007/s11419-015-0286-5.
[5] A. Helander, M. Bäckberg, P. Hultén, Y. Al-Saffar, O. Beck, Detection of new psychoactive substance use among emergency room patients: results from the Swedish STRIDA project, Forensic Sci. Int. 243 (2014) 23–29, https://doi.org/10.1016/j.forsciint.2014.02.022.
[6] S. Zaami, R. Giorgetti, S. Pichini, F. Pantano, E. Marinelli, F.P. Busardò, Synthetic cathinones related fatalities: an update, Eur. Rev. Med. Pharmacol. Sci. 22 (1) (2018) 268–274, https://doi.org/10.26355/eurrev_201801_14129.
[7] C.C. Lin, T.I. Weng, C.J. Ng, C.P. Shih, J. Hsu, Y.C. Liao, et al., Emergency department visits due to new psychoactive substances and other illicit drugs in Taiwan: preliminary results of the Taiwan Emergency Department Drug Abuse Surveillance (TEDAS) project, Clin. Toxicol. (Phila.) 60 (6) (2022) 708–715, https://doi.org/10.1080/15563650.2022.2038793.
[8] M.R. Meyer, H.H. Maurer, Current applications of high-resolution mass spectrometry in drug metabolism studies, Anal. Bioanal. Chem. 403 (5) (2012) 1221–1231, https://doi.org/10.1007/s00216-012-5824-2.
[9] C. Montesano, G. Vannutelli, A. Gregori, et al., Broad screening and identification of novel psychoactive substances in plasma by high-performance liquid chromatography-high-resolution mass spectrometry and post-run library matching, J. Chromatogr. B 1057 (2017) 88–95, https://doi.org/10.1016/j.jchromb.2017.04.037.
[10] R.J. Dinis-Oliveira, Metabolomics of drugs of abuse: a more realistic view of the toxicological complexity, Bioanalysis 6 (22) (2014) 3155–3159, https://doi.org/10.4155/bio.14.260.
[11] A. Saitman, H.D. Park, R.L. Fitzgerald, False-positive interferences of common urine drug screen immunoassays: a review, J. Anal. Toxicol. 38 (7) (2014) 387–396, https://doi.org/10.1093/jat/bku075.
[12] A. Begeman, E.J.F. Franssen, Lack of detection of new amphetamine-like drugs using conventional urinary immunoassays, Ther. Drug Monit. 40 (1) (2018) 135–139, https://doi.org/10.1097/FTD.0000000000000475.
[13] L.E. Regester, J.D. Chmiel, J.M. Holler, S.P. Vorce, B. Levine, T.Z. Bosy, Determination of designer drug cross-reactivity on five commercial immunoassay

screening kits, J. Anal. Toxicol. 39 (2) (2015) 144–151, https://doi.org/10.1093/jat/bku133.
[14] M.D. Krasowski, S. Ekins, Using cheminformatics to predict cross-reactivity of "designer drugs" to their currently available immunoassays, J Cheminform. 6 (2014) 22, https://doi.org/10.1186/1758-2946-6-22.
[15] I. Kononenko, Machine learning for medical diagnosis: history, state of the art and perspective, Artif. Intell. Med. 23 (1) (2001) 89–109, https://doi.org/10.1016/s0933-3657(01)00077-x.
[16] L. Zhang, H. Zhang, H. Ai, et al., Applications of machine learning methods in drug toxicity prediction, Curr. Top. Med. Chem. 18 (12) (2018) 987–997, https://doi.org/10.2174/1568026618666180723141939.
[17] Y.H. Hu, S.S. Wu, C.T. Tai, et al., Machine learning applications for predicting substance use disorders: a systematic review, J. Subst. Abuse Treat. 116 (2020) 108058, https://doi.org/10.1016/j.jsat.2020.108058.
[18] L.Y. Feng, J.H. Li, New psychoactive substances in Taiwan: challenges and strategies, Curr. Opin. Psychiatry 33 (4) (2020) 306–311, https://doi.org/10.1097/YCO.0000000000000604.
[19] T.I. Weng, L.Y. Chen, J.Y. Chen, P.S. Chen, H.L. Hwa, C.C. Fang, Characteristics of analytically confirmed illicit substance-using patients in the Emergency Department, J. Formos. Med. Assoc. 119 (12) (2020) 1827–1834, https://doi.org/10.1016/j.jfma.2020.01.005.
[20] W.J. Chen, S.C. Wu, W.I. Tsay, Y.T. Chen, P.C. Hsiao, Y.H. Yu, et al., Differences in prevalence, socio-behavioral correlates, and psychosocial distress between club drug and hard drug use in Taiwan: results from the 2014 National Survey of Substance Use, Int. J. Drug Policy 48 (2017) 99–107, https://doi.org/10.1016/j.drugpo.2017.07.003.
[21] D.M. Wood, F. Heyerdahl, C.B. Yates, A.M. Dines, I. Giraudon, K.E. Hovda, et al., The European Drug Emergencies Network (Euro-DEN), Clin. Toxicol. (Phila.) 52 (4) (2014) 239–241, https://doi.org/10.3109/15563650.2014.898771.
[22] B. Norgeot, G. Quer, B.K. Beaulieu-Jones, A. Torkamani, R. Dias, M. Gianfrancesco, et al., Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist, Nat. Med. 26 (9) (2020) 1320–1324, https://doi.org/10.1038/s41591-020-1041-y.
[23] Scott M. Lundberg, Su-In Lee, A unified approach to interpreting model predictions. Neural Information Processing Systems (NIPS). Oral Presentation, 2017.
[24] S. Džeroski, P. Panov, B. Ženko, Ensemble methods in machine learning. In: Meyer T (ed) Encyclopedia of Complexity and Systems Science. Springer, New York, 2009, pp 5317–5325.
[25] D. Bzdok, N. Altman, M. Krzywinski, Statistics versus machine learning, Nat. Methods 15 (4) (2018) 233–234, https://doi.org/10.1038/nmeth.4642.
[26] L. Prokhorenkova, G. Gusev, A. Vorobev, A.V. Dorogush, A. Gulin, CatBoost: unbiased boosting with categorical features, Adv. Neural Inf. Process Syst. 31 (2018) 6638–6648.
[27] G. Ke, Q. Meng, T. Finley, et al., LightGBM: a highly efficient gradient boosting decision tree, Adv. Neural Inf. Process Syst. 30 (2017) 3146–3154.
[28] R. Trevethan, Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice, Front. Public Health 5 (2017) 307, https://doi.org/10.3389/fpubh.2017.00307.
[29] P. Ranganathan, R. Aggarwal, Common pitfalls in statistical analysis: understanding the properties of diagnostic tests - Part 1, Perspect. Clin. Res. 9 (1) (2018) 40–43, https://doi.org/10.4103/picr.PICR_170_17.
[30] T.O. Ihongbe, S.W. Masho, Prevalence, correlates and patterns of heroin use among young adults in the United States, Addict. Behav. 63 (2016) 74–81, https://doi.org/10.1016/j.addbeh.2016.07.003.
[31] M. Kariisa, L. Scholl, N. Wilson, P. Seth, B. Hoots, Drug overdose deaths involving cocaine and psychostimulants with abuse potential - United States, 2003-2017, MMWR Morb. Mortal. Wkly Rep. 68 (17) (2019) 388–395, https://doi.org/10.15585/mmwr.mm6817a3.
[32] R.M. Gladden, J. O'Donnell, C.L. Mattson, P. Seth, Changes in opioid-involved overdose deaths by opioid type and presence of benzodiazepines, cocaine, and methamphetamine - 25 States, July-December 2017 to January-June 2018, MMWR Morb. Mortal. Wkly Rep. 68 (34) (2019) 737–744, https://doi.org/10.15585/mmwr.mm6834a2.
[33] J.J. Palamar, A. Le, P. Mateu-Gelabert, Not just heroin: extensive polysubstance use among US high school seniors who currently use heroin, Drug Alcohol Depend. 188 (2018) 377–384, https://doi.org/10.1016/j.drugalcdep.2018.05.001.
[34] W.M. Compton, R.J. Valentino, R.L. DuPont, Polysubstance use in the U.S. opioid crisis, Mol. Psychiatry 26 (1) (2021) 41–50, https://doi.org/10.1038/s41380-020-00949-3.
[35] Y. Dai, P.M. Musumari, H. Chen, Y. Huang, T. Techasrivichien, S.P. Suguimoto, et al., Recreational drug use, polydrug use and sexual behaviors among men who have sex with men in southwestern china: a cross-sectional study, Behav. Med. 45 (4) (2019) 314–322, https://doi.org/10.1080/08964289.2018.1538099.
[36] J. Halladay, R. Woock, H. El-Khechen, C. Munn, J. MacKillop, M. Amlung, et al., Patterns of substance use among adolescents: a systematic review, Drug Alcohol Depend. 216 (2020) 108222, https://doi.org/10.1016/j.drugalcdep.2020.108222.