Personalized Psychiatry and Neurology

Article

The Role of Family, Microsocial and Medical History in The Shaping of Trajectories of Complex Opioid and Cannabis Addiction: Results of Machine Learning Modeling

Timur Syunyakov[1,2,3*], Inara Khayredinova[1,2,4], Zarifjon Ashurov[1,2]

Citation: Syunyakov, T.; Khayredinova, I.; Ashurov, Z. The Role of Family, Microsocial and Medical History in The Shaping of Trajectories of Complex Opioid and Cannabis Ad-diction: Results of Machine Learning Modeling. *Personalized Psychiatry and Neurology* **2023**, *3 (2)*: 120-133. https://doi.org/10.52667/2712-9179-2023-3-2-120-133

Chief Editor: Nikolaj G. Neznanov, D Med Sci, Professor

Received: 31 October 2023
Accepted: 13 November 2023
Published: 15 November 2023

Publisher's Note: V.M. Bekhterev NMRC PN stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[1] Republican Specialized Scientific and Practical Medical Center for Mental Health, Tashkent, Uzbekistan;
[2] Tashkent Medical Academy, Tashkent, Uzbekistan;
[3] Samara State Medical University, 443099, Samara, Russia;
[4] Center for Development of Professional Qualification of Medical Workers, Tashkent, Uzbekistan

* Correspondence: sjunja@gmail.com (T.S.)

Abstract: Introduction: The widespread misuse of opioids and cannabis is a notable global public health concern. The substantial public health concern due to the misuse of opioids and cannabis, individually and concurrently, is associated with vast societal implications. Identification of risk factors for developing misuse of these substances is of utmost importance. This study aims at developing a machine learning-based model to classify groups of opioid or cannabis dependents using family, microsocial, and medical history variables, and to identify the most significant variables associated with each group. **Methods:** This naturalistic observational non-interventional study enrolled adult patients diagnosed with opioid use disorder, cannabis use disorder, or a combination of both. Machine learning models, including Stacking, Logistic Regression, Gradient Boosting, k-Nearest Neighbors (kNN), Naive Bayes, Support Vector Machines (SVM), Random Forest, and Decision Tree, were used to classify patients and predict their risk factors based on various personal history variables. **Results:** The patient groups showed significant differences in their working fields, marital status before and after the formation of drug addiction, substance misuse in relatives, family type, parent-child relationships, and birth order. They also differed significantly in fleeing from home and personality types. Machine learning models provided high classification accuracy across all substance dependence groups, particularly for the cannabis group (>90% accuracy). Significant differences were found among the complex misuse group, where individuals faced severe psychosocial issues originating from the familial environment, such as a history of fleeing home, coming from a single-parent family, and dominant parent-child relationships. **Discussion:** The methods used in this study provided robust and reliable assessments of the models' predictive performances. The results pointed to significant differences in familial and developmental factors between the three dependence groups. The complex dependence group showed more severe psychosocial issues originating from the family environment. This group also revealed a specific sequence of life events and conditions predictive of complex dependence. These findings highlight the importance of interventions that address risk factors across various life stages and domains. **Conclusion:** Early identification of high-risk individuals and understanding the risk factors can inform the development of effective interventions at both individual and societal levels, ultimately aiming at mitigating dependence risks and improving overall well-being. Further research with longitudinal designs and diverse populations are needed to increase our understanding of trajectory of addiction formation in order to deliver effective interventions for individuals at risk.

Keywords: *opioids; cannabis; complex dependence; machine learning; medical history.*

Introduction

Opioid and cannabis addiction is an important public health problem, has profound health and social consequences, and is associated with significant morbidity and mortality

[1-4]. These disorders are characterized by a high burden on healthcare system due to the associated costs of treatment and have a large impact on work productivity and social relationships.

Complex dependence on cannabis and opioids is of particular concern as it potentially can combine the risks associated with each disorder separately and could potentially lead to more severe social and health consequences [5]. Moreover, the two disorders appear to be tightly interrelated. Williams et al. (2020) found that cannabis use was associated with an increased risk of developing opioid use disorder, particularly among those who started using cannabis at an earlier age, suggesting that cannabis exposure may prepare the brain for the development of opioid dependence, making people more susceptible [6]. A longitudinal analysis by Lake at al. (2019) found that the risk of developing opioid use disorder was significantly higher in people who used both cannabis and opioids for chronic pain compared to those who used only one of these substances [7]. On the other hand, Liang et al. (2019) found that medical cannabis was associated with a reduced risk of opioid addiction, especially among people using it to treat chronic pain [8]. Nevertheless, in another study of opioids and cannabis use for chronic pain individuals who co-use both substances are at an increased risk of substance misuse, mental health problems, and worse pain experiences and associated with a range of negative outcomes, including increased risk of overdose and opioid-related disorders [9]. This suggests that the complex use of cannabis and opioids can potentially have a synergistic effect that exacerbates the negative effects of each substance and may therefore be particularly detrimental, increasing the likelihood of developing dependence on each substance, which may require a more complex treatment and recovery regimen.

Therefore, understanding the risk factors for developing opioid, cannabis, and complex misuse is crucial. These factors may include, but are not limited to, family history, social environment, and broader sociocultural influences.

A large survey of adolescents and young adults in the United States identified various risk factors for opioid abuse, including mood and anxiety disorders, male gender, educational level, and a history of substance abuse [10]. Similarly, risk factors for cannabis misuse include younger age, male gender, and co-occurring alcohol use [8]. Knowledge of these risk factors will allow health professionals and decision makers to identify risk factors and high-risk populations to facilitate the development and implementation of effective primary prevention strategies. Moreover, it will allow risk stratification in the population and tailor primary prevention to the needs of each risk group, thereby optimizing resource allocation and improving the effectiveness of interventions [10].

Considering this, the aims of this study to create machine learning-based model to classify groups of patients who have opioid or cannabis dependence alone, or both, using family, microsocial, and medical history variables, and to find which variables have most significant association with each group.

Materials and Methods

Patients

All consecutive adult (at least 18 years) patients who admitted to the Republican Specialized Scientific and Practical Center of Narcology diagnosed with opioid (Tramadol) dependence (F11.2) alone, cannabis (except synthetic cannabinoids) dependence (F12.2) alone, or combination of both according to ICD-10 criteria were enrolled in this naturalistic observational study. During a thorough clinical interview the comprehensive set of information on patients' demographic profiles, relationships, family backgrounds, developmental histories, behaviors, psychology, and medical histories was collected as following:

1. Demographic information (age, disease duration).

2. Socioeconomic status: education level, working field.
3. Relationship factors: marital status before and after addiction formation, parents-child relationships, family type.
4. Family medical history and history of substance use: psychiatric disorders in relatives, alcohol consumption of parents, substance misuse in father and relatives, alcohol or substance misuse in other relatives.
5. Birth and developmental history: order of birth, antenatal and intranatal pathology, perinatal pathology.
6. Behavioral and lifestyle factors: school discipline, fleeing from home, participation in asocial/criminal groups, criminal history.
7. Psychological attributes: deviant behavior, personality type.

In every case the information was cross-checked from the patients' relatives.

Statistical Analysis

For continuous variables (age), a mean value and standard deviation (SD) was reported, and between-group differences were tested with one-way ANOVA. For categorical variables we reported absolute and relative frequencies, such as disease duration, education, working field, marital status, family type, etc., while to assess the independence of distribution of frequencies between the three groups for these variables a chi-square ($\chi 2$) test was used.

In this study, to understand factors influencing different dependence groups we used several machine learning (ML) models to predict outcomes in three distinct groups: cannabis, opioid, and complex groups. The models utilized were Stacking, Logistic Regression, Gradient Boosting, k-Nearest Neighbors (kNN), Naive Bayes, Support Vector Machines (SVM), Random Forest, and Decision Tree.

To fit and evaluate these models, we partitioned our data into training and test datasets, with 66% of the data used for training each model and the remaining 33% used to test the model's predictions. This allows us to assess the predictive performance of each model and ensure it generalizes well to unseen data.

Additionally, instead of using traditional cross-validation methods, we employed a repeated random sub-sampling validation approach. This involved randomly dividing the dataset into a training set and a test set 20 times, fitting the model to each training set, and then evaluating it on the corresponding test set.

The performance of each model was evaluated using a suite of metrics, including Area Under the Receiver Operating Characteristic Curve (AUC), Classification Accuracy (CA), F1-score, precision, recall, and Matthews Correlation Coefficient (MCC). These metrics provide insights into the model's ability to discriminate between classes, its overall accuracy, its balance between precision and recall, and the correlation between observed and predicted binary classifications, respectively. Best model was selected based on the classification accuracy in every group with subsequent construction of beeswarm plots to explain impact of variables in respective classification category. Statistical analysis was performed using programming in Python via pandas, scipy, scikit-learn, and SHAP libraires [11-15].

Results

Overall, 129 patients (all males) were included in the study: n=32 in the cannabis group, n=44 in the opioid group, and n=53 in the complex group. Their general characteristics and information on medical history and family and microsocial variables used in the predictive models in the groups of patients with opioid, cannabis and complex dependence are described in Table 1.

The groups showed significant differences in their working fields ($p = 0.007$). The cannabis group had a high proportion working in transport and logistics (40.62%). The opioid group had the highest proportion in the marketing field (36.36%), and individuals

in the complex dependence group were mostly working in the small business and retail trade sector (43.40%). Marital status before the formation of drug addiction showed significant differences between groups (p = 0.014).

Notably, more people in the complex dependence group were married before drug addiction formed. There was also a significant shift in marital status after the formation of drug addiction (p = 0.038), with an increase in the proportion of married individuals in all groups. Substance misuse in relatives was significantly different between groups (p < 0.001). Most notably, the complex dependence group had a much higher proportion of individuals (62.26%) who have other relatives misusing substances. There was a significant difference between groups in terms of family type (p = 0.002).

The largest proportion (73.58%) of the complex dependence group came from a single-parent family. Significant differences were found in parents-child relationships among the groups (p = 0.001). Specifically, the complex dependence group had the highest proportion of dominant hyperprotection (60.38%). The order of birth showed significant differences between the groups (p < 0.001). Most notably, the complex dependence group had a higher percentage of individuals who are an only child. There was a significant difference in the 'Flee from home' category (p = 0.002). The complex dependence group had the highest percentage of individuals who had fled from home.

There was a significant difference in personality types among the groups (p = 0.001). The complex dependence group had a higher proportion of individuals with unstable personality types.

Table 1. General characteristics and information on medical history and family and microsocial variables used in the predictive models in the groups of patients with opioid, cannabis and complex dependence.

Variable	Levels	Cannabis dependence group (n=32)	Opioid dependence group (n=44)	Complex dependence group (n=53)	Test statistic
Age, Mean (SD)		31.625 (7.106)	30.068 (6.760)	31.113 (6.947)	F = 0.518, df = 2/126, p = 0.597
Disease duration, n (%)	1-3 years	6 (18.75%)	12 (27.27%)	13 (24.53%)	χ^2 = 1.055, df = 4, p = 0.901
	3-5 years	17 (53.12%)	23 (52.27%)	28 (52.83%)	
	5-8 years	9 (28.12%)	9 (20.45%)	12 (22.64%)	
Education, n (%)	University	7 (21.88%)	9 (20.45%)	8 (15.09%)	χ^2 = 2.740, df = 4, p = 0.602
	High school	12 (37.50%)	21 (47.73%)	20 (37.74%)	
	College	13 (40.62%)	14 (31.82%)	25 (47.17%)	
Working field, n (%)	Education	0 (0.00%)	7 (15.91%)	2 (3.77%)	χ^2 = 20.957, df = 8, p = 0.007
	Transport and logistics	13 (40.62%)	11 (25.00%)	18 (33.96%)	
	Civil service	3 (9.38%)	1 (2.27%)	3 (5.66%)	
	Marketing	7 (21.88%)	16 (36.36%)	7 (13.21%)	
	Small business and retail trade sector	9 (28.12%)	9 (20.45%)	23 (43.40%)	
Marital status before the formation of drug addiction, n (%)	Married	13 (40.62%)	18 (40.91%)	29 (54.72%)	χ^2 = 12.429, df = 4, p = 0.014
	Divorced	12 (37.50%)	5 (11.36%)	9 (16.98%)	
	Single	7 (21.88%)	21 (47.73%)	15 (28.30%)	

Marital status after the formation of drug addiction, n (%)	Married	17 (53.12%)	24 (54.55%)	40 (75.47%)	$\chi^2 = 10.151$, df = 4, p = 0.038
	Divorsed	13 (40.62%)	12 (27.27%)	10 (18.87%)	
	Single	2 (6.25%)	8 (18.18%)	3 (5.66%)	
Alcohol in one parent, n (%)	No	20 (62.50%)	28 (63.64%)	36 (67.92%)	$\chi^2 = 0.323$, df = 2, p = 0.851
	Yes	12 (37.50%)	16 (36.36%)	17 (32.08%)	
Alcohol in both parents, n (%)	No	27 (84.38%)	36 (81.82%)	46 (86.79%)	$\chi^2 = 0.455$, df = 2, p = 0.797
	Yes	5 (15.62%)	8 (18.18%)	7 (13.21%)	
Alcohol or substance misuse in other relatives, n (%)	No	25 (78.12%)	32 (72.73%)	20 (37.74%)	$\chi^2 = 18.245$, df = 2, p < 0.001
	Yes	7 (21.88%)	12 (27.27%)	33 (62.26%)	
Alcohol misuse in relatives, n (%)	No	32 (100.00%)	44 (100.00%)	50 (94.34%)	$\chi^2 = 4.404$, df = 2, p = 0.111
	Yes	0 (0.00%)	0 (0.00%)	3 (5.66%)	
Alcohol misuse in 2-3 relatives, n (%)	No	32 (100.00%)	44 (100.00%)	53 (100.00%)	$\chi^2 < 0.001$, df = 0, p = 1.000
Substance misuse in father, n (%)	No	28 (87.50%)	40 (90.91%)	45 (84.91%)	$\chi^2 = 0.798$, df = 2, p = 0.671
	Yes	4 (12.50%)	4 (9.09%)	8 (15.09%)	
Substance misuse in 1 relative, n (%)	No	29 (90.62%)	40 (90.91%)	51 (96.23%)	$\chi^2 = 1.425$, df = 2, p = 0.491
	Yes	3 (9.38%)	4 (9.09%)	2 (3.77%)	
Substance misuse in 2-3 relatives, n (%)	No	29 (90.62%)	39 (88.64%)	34 (64.15%)	$\chi^2 = 12.143$, df = 2, p = 0.002
	Yes	3 (9.38%)	5 (11.36%)	19 (35.85%)	
Psychiatric disorders in relatives, n (%)	No	29 (90.62%)	42 (95.45%)	50 (94.34%)	$\chi^2 = 0.788$, df = 2, p = 0.674
	Yes	3 (9.38%)	2 (4.55%)	3 (5.66%)	
Antenatal and intranatal pathology, n (%)	Yes	0 (0.00%)	0 (0.00%)	4 (7.55%)	$\chi^2 = 5.919$, df = 2, p = 0.052
	No	32 (100.00%)	44 (100.00%)	49 (92.45%)	
Perinatal pathology, n (%)	Asphixia	0 (0.00%)	1 (2.27%)	3 (5.66%)	$\chi^2 = 3.638$, df = 6, p = 0.725
	Other	1 (3.12%)	0 (0.00%)	1 (1.89%)	
	Prematurity	1 (3.12%)	1 (2.27%)	1 (1.89%)	
	No	30 (93.75%)	42 (95.45%)	48 (90.57%)	
Family type, n (%)	One parent	11 (34.38%)	17 (38.64%)	39 (73.58%)	$\chi^2 = 17.266$, df = 4, p = 0.002
	Legal guardian	4 (12.50%)	4 (9.09%)	2 (3.77%)	
	Complete family	17 (53.12%)	23 (52.27%)	12 (22.64%)	
Parents-child relationships, n (%)	Hypoprotection	14 (43.75%)	10 (22.73%)	7 (13.21%)	$\chi^2 = 31.295$, df = 10, p = 0.001
	Dominant hyperprotection	7 (21.88%)	11 (25.00%)	32 (60.38%)	
	Abusive relationships	0 (0.00%)	2 (4.55%)	2 (3.77%)	
	Increased moral responsibility	2 (6.25%)	0 (0.00%)	1 (1.89%)	
	Coercive hyperprotection	5 (15.62%)	17 (38.64%)	9 (16.98%)	
	Emotional rejection	4 (12.50%)	4 (9.09%)	2 (3.77%)	

Order of birth, n (%)	Only child	8 (25.00%)	12 (27.27%)	27 (50.94%)	$\chi^2 = 28.284$, df = 4, p
	Middle or youngest	17 (53.12%)	23 (52.27%)	4 (7.55%)	< 0.001
	Elder	7 (21.88%)	9 (20.45%)	22 (41.51%)	
School discipline, n	No	20 (62.50%)	28 (63.64%)	22 (41.51%)	$\chi^2 = 5.906$, df = 2, p =
(%)	Yes	12 (37.50%)	16 (36.36%)	31 (58.49%)	0.052
Flee from home, n	No	23 (71.88%)	36 (81.82%)	26 (49.06%)	$\chi^2 = 12.159$, df = 2, p
(%)	Yes	9 (28.12%)	8 (18.18%)	27 (50.94%)	= 0.002
Asocial group	No	21 (65.62%)	32 (72.73%)	27 (50.94%)	$\chi^2 = 5.078$, df = 2, p =
participation, n (%)	Yes	11 (34.38%)	12 (27.27%)	26 (49.06%)	0.079
Criminal group	No	29 (90.62%)	42 (95.45%)	48 (90.57%)	$\chi^2 = 0.960$, df = 2, p =
participation, n (%)	Yes	3 (9.38%)	2 (4.55%)	5 (9.43%)	0.619
Criminal history, n	No	30 (93.75%)	44 (100.00%)	50 (94.34%)	$\chi^2 = 2.711$, df = 2, p =
(%)	Yes	2 (6.25%)	0 (0.00%)	3 (5.66%)	0.258
No deviant	No	15 (46.88%)	20 (45.45%)	37 (69.81%)	$\chi^2 = 7.162$, df = 2, p =
bechavior, n (%)	Yes	17 (53.12%)	24 (54.55%)	16 (30.19%)	0.028
Personality, n (%)	Asthenic	3 (9.38%)	0 (0.00%)	1 (1.89%)	$\chi^2 = 39.324$, df = 16,
	Histrionic	3 (9.38%)	10 (22.73%)	4 (7.55%)	p = 0.001
	Conforming	4 (12.50%)	9 (20.45%)	10 (18.87%)	
	Unstable	4 (12.50%)	13 (29.55%)	27 (50.94%)	
	Paranoid	0 (0.00%)	1 (2.27%)	0 (0.00%)	
	Psychasthenic	7 (21.88%)	3 (6.82%)	6 (11.32%)	
	Cycloid	1 (3.12%)	2 (4.55%)	3 (5.66%)	
	Schizoid	0 (0.00%)	1 (2.27%)	0 (0.00%)	
	Epileptoid	10 (31.25%)	5 (11.36%)	2 (3.77%)	

Table 2. Comparison of predictive models that explain classification into groups based on the medical history.

Model	AUC	CA	F1	Precision	Recall	MCC
Complex group						
Stack	0.917	0.843	0.789	0.814	0.766	0.665
Logistic Regression	0.897	0.825	0.769	0.776	0.763	0.629
Gradient Boosting	0.896	0.823	0.767	0.772	0.763	0.624
kNN	0.898	0.817	0.770	0.743	0.798	0.619
Naive Bayes	0.902	0.807	0.723	0.801	0.659	0.583
SVM	0.912	0.800	0.765	0.695	0.852	0.604
Random Forest	0.856	0.775	0.711	0.700	0.721	0.527
Tree	0.790	0.762	0.689	0.692	0.685	0.497
Cannabis group						
Stack	0.986	0.935	0.879	0.870	0.889	0.835
Gradient Boosting	0.963	0.933	0.873	0.879	0.868	0.828
kNN	0.982	0.928	0.863	0.880	0.846	0.815
SVM	0.983	0.928	0.851	0.952	0.769	0.813

Naive Bayes	0.982	0.924	0.863	0.827	0.902	0.812
Logistic Regression	0.970	0.917	0.843	0.848	0.838	0.787
Random Forest	0.963	0.910	0.834	0.819	0.850	0.773
Tree	0.901	0.887	0.797	0.767	0.829	0.720
			Opioid group			
Stack	0.900	0.822	0.752	0.735	0.770	0.613
SVM	0.908	0.810	0.716	0.755	0.680	0.576
Gradient Boosting	0.892	0.806	0.726	0.718	0.735	0.576
Logistic Regression	0.863	0.792	0.708	0.698	0.718	0.547
Naive Bayes	0.885	0.792	0.721	0.681	0.767	0.559
kNN	0.865	0.789	0.691	0.710	0.673	0.531
Random Forest	0.838	0.776	0.671	0.693	0.650	0.502
Tree	0.790	0.757	0.645	0.662	0.628	0.460

The predictive performance of various machine learning models for classification into three dependence groups (complex, cannabis, opioid) based on medical history data is presented in the Table 2. Across datasets, Stacking generally achieved the best performance, with AUC ranging from 0.900-0.986, followed by Gradient Boosting and kNN. Naive Bayes and Decision Tree tended to underperform. Models performed best on the cannabis group (AUC 0.901-0.986) compared to complex (AUC 0.790-0.917) and opioid groups (AUC 0.838-0.908). For the Gradient Boosting model classification accuracy for the complex group was 82.3%, for cannabis group – 93.3%, and for opioid group – 80.6%.

The beeswarm plots for the predictors of complex, cannabis, and opioids groups can be seen on the Fig. 1-3. In the cannabis dependence group, substance use patterns were strongly influenced by family members, as illustrated by a positive association with following family members' substance use examples. Furthermore, an earlier age of initiating alcohol use from 17 years old was also positively associated, suggesting an early environmental influence on substance use patterns within this group. In contrast, for the opioid dependence group, positive family dynamics and early life stability appeared to be significantly associated. Features such as the lack of history of fleeing from home, absence of negative substance-use influences from family members, and non-single parent family type were leading factors in distinguishing this group. Further, being a younger sibling in the family also emerged as a significant characteristic. For the complex dependence group, a variety of challenging circumstances were shown to be strong predictors. Having a history of fleeing from home, being from a single-parent family, and experiencing dominant hyperprotection in parent-child relationships were positively associated. Furthermore, any age of first cannabis use and being the first-born child in the family emerged as key predictors.

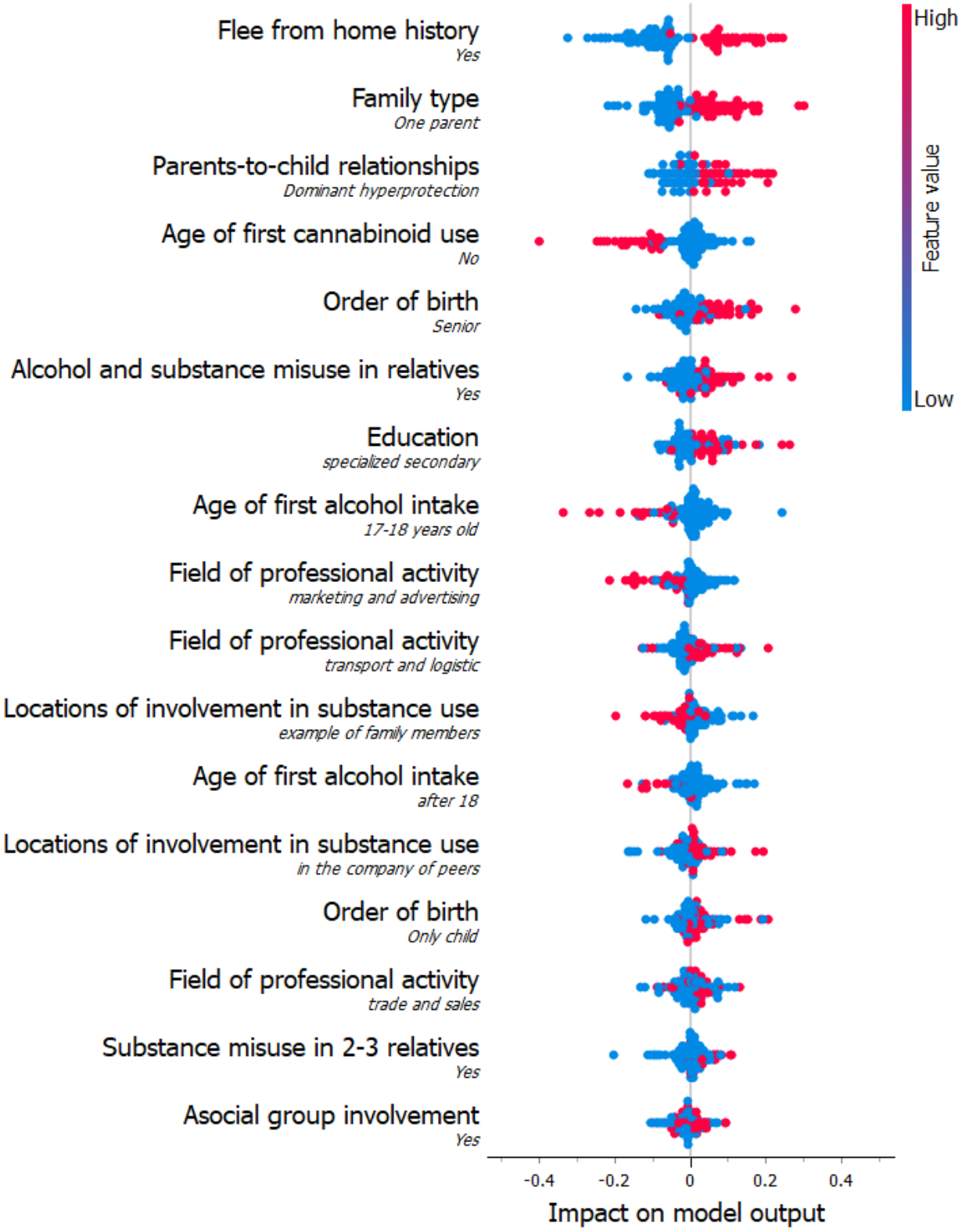

Figure 1. Top factors explaining the Gradient Boosting model for the complex dependence group.

Figure 2. Top factors explaining the Gradient Boosting model for the cannabis dependence group.

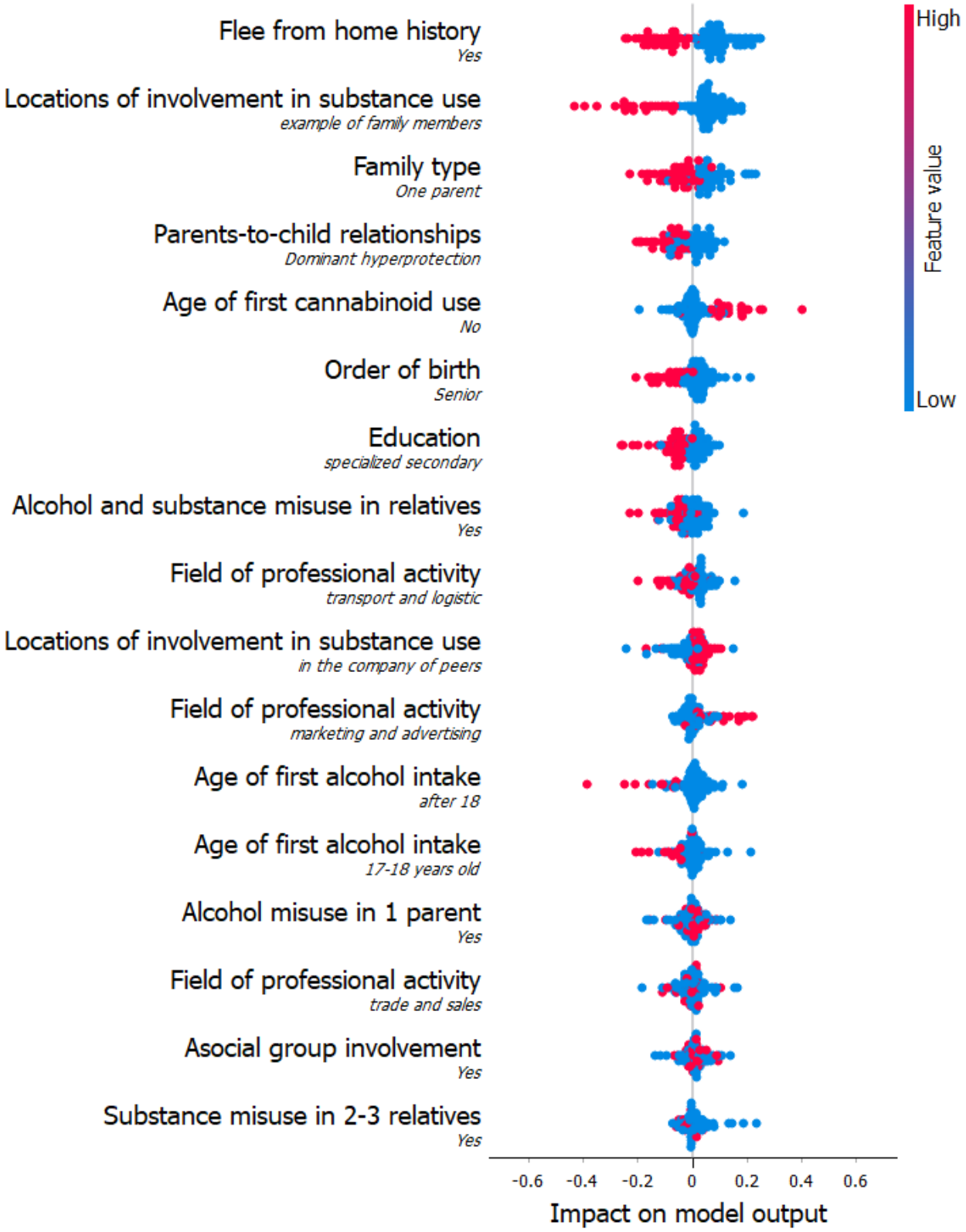

Figure 3. Top factors explaining the Gradient Boosting model for the opioid dependence group.

Discussion

In this study we used a diverse range of ML models to get insights on what factors from personal history, including family, social, and medical variables could allow us to predict, if patient will develop dependence from opioid (tramadol) or cannabis, or both substances. We found that all models provided high classification accuracy (>80%) in distinguishing each group of patients with particularly high accuracy (>90%) for cannabis group.

The results suggest differences in important familial and developmental factors between the three dependence groups. For the opioid dependence group, factors such as history of fleeing from home, substance use involving family members, coming from a one-parent family, and experiences of dominant hyperprotection in parenting were negatively associated. In the cannabis dependence group, substance use involving family members and earlier alcohol use from 17-18 years old showed positive associations, though relationships were uncertain for most factors.

The complex dependence group displayed positive associations with fleeing from home, coming from a one-parent family, experiences of dominant hyperprotection parenting, and earlier family exposure to substance and alcohol use. Additionally, more factors in this group demonstrated clear positive or negative impacts compared to the cannabis dependence group. These results are in line with other studies, that observed that a history of fleeing from home, substance use involving family members, coming from a one-parent family, and parent-child relationships as potential risk factors for substance abuse [16, 17].

This suggests those in the complex dependence group faced more severe psychosocial issues originating from the family environment, as evidenced by the higher rankings and clearer directions of impactful factors. The results point to important differences in familial and developmental characteristics between the three clinical profiles. Experiences such as problematic parenting, substance exposure in the home, and family structure appear to have distinguished the complex dependence group, implicating greater socioenvironmental adversity in this population. These findings warrant further exploration to better understand heterogeneity in substance dependence phenotypes.

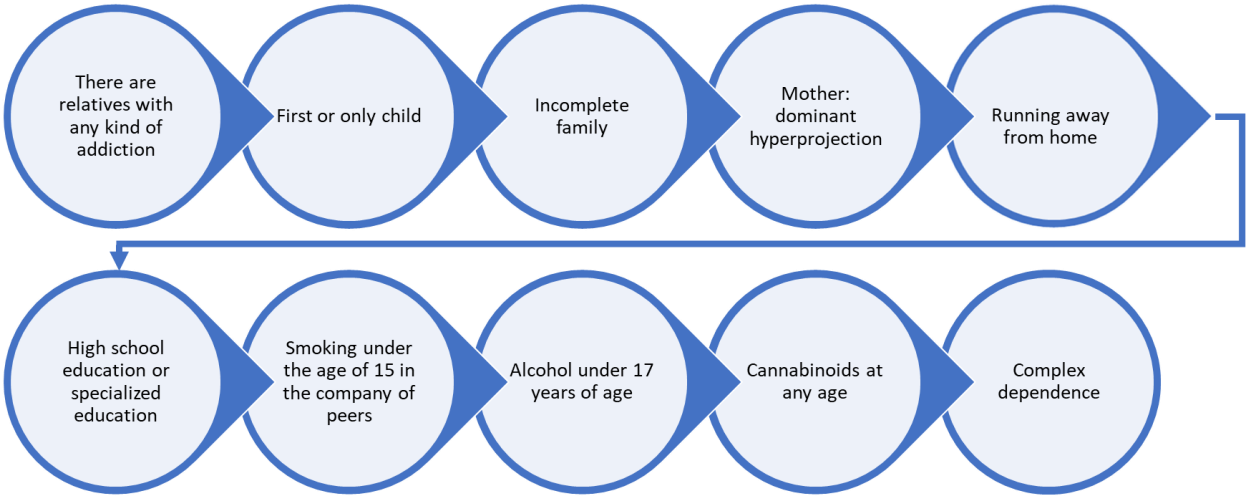

Figure 4. Trajectory of life events to predict complex dependence.

Based on our results we can hypothesize a specific sequence of life events and conditions emerged as predictive of complex dependence (Fig. 4). This trajectory begins with a familial predisposition to addiction, highlighting the influence of genetic and

environmental predispositions on the development of complex dependence. The individual's birth order and family structure also play significant roles. Being the first-born or an only child in an incomplete family, particularly with a dominant mother, contributes to the susceptibility to complex dependence. This scenario underlines the potential impacts of early life family dynamics and parental styles on the risk of developing substance use disorders. Subsequent life events, notably running away from home, early onset of smoking (under the age of 15) in the company of peers, and initiating alcohol (under 17 years of age) or cannabis use at any age during adolescence add further layers of risk. The initiation of these substance use behaviors at such a young age, often within social contexts, underscores the role of peer influence and potentially indicative of attempts to cope with adversity or stress. Finally, the individual's level of education — not reaching beyond high school or specialized education — emerged as a predictive factor. This suggests the potential influence of socio-economic factors and the potential repercussions of early onset substance misuse on educational attainment. This sequence of life events and conditions provides a comprehensive view of the multifaceted and complex nature of the risk trajectory towards complex dependence and expands the existing literature on the role of early stress, family environment and other predisposition factors in the formation of addiction to multiple substances [18]. Finally, our results underscore the necessity of interventions that address these risk factors across various stages of life and domains.

The obtained data is of interest in the context of identifying risk factors and determining high-risk groups that can be targeted for primary prevention by healthcare, education, social services, and supervisory bodies. This information is crucial for developing and implementing effective interventions at individual and societal levels, ultimately aimed at mitigating the risk of dependence and improving overall well-being. Through identifying the high-risk groups, appropriate resources and services can be targeted more effectively, ensuring that those most at risk receive the necessary support and intervention.

Conclusions

Our study identified key risk factors for developing opioid, cannabis, and complex addiction. These insights provide valuable information that can guide the development of targeted preventive strategies. Importantly, early identification of at-risk individuals or groups enables the implementation of primary prevention strategies, which have proven effective in reducing the risk of substance misuse. Our results demonstrate distinct trajectories of different types of substance misuse.

The implications of these findings may be crucial for healthcare, education, social services, and supervisory bodies, since it may help in the identification of high-risk individuals while an understanding the pathways leading to substance misuse can enable these bodies to develop targeted prevention strategies. Such an approach expected to reduce societal and economic burdens associated with substance misuse.

More research employing longitudinal designs with larger, diverse populations are needed to track the progression of substance misuse over time and to get more precise predictive models. This would allow us to capture the dynamic interplay of risk factors and their influence on substance misuse. Such efforts could enhance our ability to design and deliver effective interventions for those at risk of chemical addiction.

Limitations

Only men were included in the study, which limits the ability to generalize the findings to a larger population. The sample size was small, which may affect the statistical power of the study. In addition, the study did not include a control group of healthy individuals, which could have provided a comparative baseline of the identified risk factors,

which would have made the predictive models more accurate and increased their generalizability.

Author Contributions: I.K. and Z.A. formulated the primary idea, I.K. designed the study, T.S. did data analysis and wrote the first draft of the manuscript, I.K., T.S., and Z.A. contributed to the detailed revision, as well as approved the final version of the article for its submission. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Faller, J.; Le, LK-D.; Chatterton, M.L.; Perez, J.K.; Chiotelis, O.; Tran, H.N.Q. et al. A systematic review of economic evaluations for opioid misuse, cannabis and illicit drug use prevention. *BJPsych Open*. **2023**; 9(5). doi: 10.1192/bjo.2023.515.
2. Lewer, D.; Freer, J.; King, E.; Larney, S.; Degenhardt, L.; Tweed, E.J. et al. Frequency of health‐care utilization by adults who use illicit drugs: a systematic review and meta‐analysis. *Addiction.* **2020**; 115(6): 1011-1023. doi: 10.1111/add.14892.
3. Aldridge, R.W.; Story, A.; Hwang, S.W.; Nordentoft, M.; Luchenski, S.A.; Hartwell, G. et al. Morbidity and mortality in homeless individuals, prisoners, sex workers, and individuals with substance use disorders in high-income countries: a systematic review and meta-analysis. *The Lancet*. **2018**; 391(10117) :241-50.
4. Lewer, D.; Tweed, E.J.; Aldridge, R.W.; Morley, K.I. Causes of hospital admission and mortality among 6683 people who use heroin: A cohort study comparing relative and absolute risks. *Drug and alcohol dependence*. **2019**; 204: 107525.
5. Reisfield, G.M. Medical cannabis and chronic opioid therapy. *J Pain Palliat Care Pharmacother*. **2010**; 24(4): 356-61. doi: 10.3109/15360288.2010.519431.
6. Williams, A.R. Cannabis as a Gateway Drug for Opioid Use Disorder. *The Journal of law, medicine & ethics : a journal of the American Society of Law, Medicine & Ethics*. **2020**; 48(2): 268-74. doi: 10.1177/1073110520935338.
7. Lake, S.; Walsh, Z.; Kerr, T.; Cooper, Z.D.; Buxton, J.; Wood, E. et al. Frequency of cannabis and illicit opioid use among people who use drugs and report chronic pain: A longitudinal analysis. *PLoS medicine.* **2019**; 16(11): e1002967. doi: 10.1371/journal.pmed.1002967.
8. Liang, D.; Wallace, M.S.; Shi, Y. Medical and non-medical cannabis use and risk of prescription opioid use disorder: Findings from propensity score matching. *Drug Alcohol Rev*. **2019**; 38(6): 597-605. doi: 10.1111/dar.12964.
9. Rogers, A.H.; Bakhshaie, J.; Buckner, J.D.; Orr, M.F.; Paulus, D.J.; Ditre, J.W. et al. Opioid and Cannabis Co-Use among Adults With Chronic Pain: Relations to Substance Misuse, Mental Health, and Pain Experience. *J Addict Med.* **2019**; 13(4): 287-294. doi: 10.1097/ADM.0000000000000493.
10. Hudgins, J.D.; Porter, J.J.; Monuteaux, M.C.; Bourgeois, F.T. Prescription opioid use and misuse among adolescents and young adults in the United States: A national survey study. *PLoS medicine*. **2019**; 16(11): e1002922. doi: 10.1371/journal.pmed.1002922.
11. Reback, J.; McKinney, W.; Van Den Bossche, J.; Augspurger, T.; Cloud, P.; Klein, A. et al. pandas-dev/pandas: Pandas 1.0. 5. *Zenodo*. **2020**.
12. McKinney, W. Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference: Austin, TX; **2010**, 51-56.
13. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods*. **2020**; 17(3): 261-272.
14. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O. et al. Scikit-learn: Machine learning in

Python. *the Journal of machine Learning research*. **2011**; 12: 2825-30.

15. Lundberg, S.M.; Lee, S-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems*. **2017**, 4765-4774.
16. Merikangas, K.R.; Dierker, L.; Fenton, B. Familial factors and substance abuse: Implications for prevention. *Drug abuse prevention through family interventions*. **1998**, 12-41.
17. Whitesell, M.; Bachand, A.; Peel, J.; Brown, M. Familial, Social, and Individual Factors Contributing to Risk for Adolescent Substance Use. *Journal of Addiction*. **2013**; 2013: 1-9. doi: 10.1155/2013/579310.
18. Konkolÿ Thege, B.; Horwood, L.; Slater, L.; Tan, M.C.; Hodgins, D.C.; Wild, T.C. Relationship between interpersonal trauma exposure and addictive behaviors: a systematic review. *BMC psychiatry*. **2017**; 17: 1-17.